

An Empirical Analysis of the Implicit Cognitions in Voice Hearing using the Implicit Relational Assessment Procedure



Thesis submitted to the Department of Psychology, Faculty of Science and Engineering, in fulfilment of the requirements for the degree of Doctor of Philosophy, National University of Ireland Maynooth.

Ciara McEnteggart (B.Sc.)

June 2015

Research Supervisor: Dr. Yvonne Barnes-Holmes

Head of Department: Dr. Andrew Coogan

Table of Contents

Acknowledgements	iii
List of Tables	iv
List of Figures.....	v
List of Appendices.....	viii
Abstract.....	ix
Chapter 1. <i>General introduction</i>	1
Chapter 2. <i>Experiments 1 and 2: Assessing stigmatisation and locus of control in psychological suffering with a non-clinical sample</i>	21
Chapter 3. <i>Experiment 3: Assessing evaluations of voice hearing in a non-voice hearing control sample</i>	52
Chapter 4. <i>Experiment 4: Assessing the effects of a voice hearing simulation on IRAP effects in a non-voice hearing control sample</i>	65
Chapter 5. <i>Experiment 5: Assessing implicit perceptions of the normality of voice hearing in non-clinical voice hearers and controls</i>	78
Chapter 6. <i>Experiment 6: Assessing the acceptance of self and others hearing voices in non-clinical voice hearers and controls</i>	92
Chapter 7. <i>Experiment 7: Assessing fear of voices in clinical and non-clinical voice hearers and controls</i>	105

Chapter 8. <i>Experiment 8: Assessing the valence and acceptance of voices in clinical and non-clinical voice hearers and controls</i>	118
Chapter 9. <i>Systematic comparisons across experiments 1-8</i>	143
Chapter 10. <i>General discussion</i>	158
References	168
Appendices	185

Acknowledgements

A big thank you to all of the participants, without whom this research wouldn't have been possible.

Yvonne – thank you so much for the opportunities you have given me over the past four years. Not only have I been lucky enough to learn from your unparalleled insight, I have made a friend. I am forever grateful.

A big thank you also to Dermot for the many valuable contributions to this work.

Thanks to all the team in The Netherlands, especially Jos, for helping us to establish such a vital research collaboration, and for your valued co-supervision throughout the thesis.

Thanks to my friends and family for your continued support (and for not asking me what I *actually* do!).

And to Noel, your definition of RFT will always make me laugh -- thanks for just being you!

List of Tables

Table 1: Stimuli and Response Options of the three IRAPs employed in Experiment 1

Table 2: Explicit Data Summary for Experiment 1

Table 3: Stimuli and Response Options of the two IRAPs employed in Experiment

Table 4: Explicit Data Summary for Experiment 2

Table 5: Stimuli and Response Options of the Evaluation IRAP employed in Experiment 3

Table 6: Explicit Data Summary for Experiment 3

Table 7: Stimuli and Response Options of the Fear IRAP employed in Experiment 4

Table 8: Explicit Data Summary for Experiment 4

Table 9: Stimuli and Response Options of the Normality IRAP employed in Experiment 5

Table 10: Explicit Data Summary for Experiment 5

Table 11: Stimuli and Response Options of the two IRAPs employed in Experiment 6

Table 12: Explicit Data Summary for Experiment 6

Table 13: Explicit Data Summary for Experiment 7

Table 14: Stimuli and Response Options in the English and Dutch of the two IRAPs
employed in Experiment 8

Table 15: Explicit Data Summary for Experiment 8

Table 16: Correlation table outlining significant correlations between Acceptance IRAP
trial-types and the explicit measures in Experiment 8

List of Figures

Figure 1: Schematic representation of the four trial-types presented in the Depression IRAP in Experiment 1

Figure 2: Mean D_{IRAP} scores for the three IRAPs in Experiment 1

Figure 3: Mean D_{IRAP} scores for the two DASS distress groups on the Depression IRAP in Experiment 1

Figure 4: Mean D_{IRAP} scores for the three DASS distress groups on the Anxiety IRAP in Experiment 1

Figure 5: Mean D_{IRAP} scores for the three DASS distress groups on the Mental Illness IRAP in Experiment 1

Figure 6: Schematic representation of the four trial-types presented in the Illness IRAP (left) and the Weakness IRAP (right) in Experiment 2

Figure 7: Mean D_{IRAP} scores for the two IRAPs in Experiment 2

Figure 8: Mean D_{IRAP} scores for the three DASS distress groups on the Illness IRAP in Experiment 2

Figure 9: Mean D_{IRAP} scores for the three DASS distress groups on the Weakness IRAP in Experiment 2

Figure 10: Schematic representation of the four trial-types presented in the Evaluation IRAP in Experiment 3

Figure 11: Mean D_{IRAP} scores for the three groups on the Evaluation IRAP in Experiment 3

Figure 12: Mean D_{IRAP} scores for the high and low CAPE groups in the Evaluation IRAP in Experiment 3

Figure 13: The four stages in Experiment 4

Figure 14: Schematic representation of the four trial-types presented in the Fear IRAP in Experiment 4

Figure 15: Mean D_{IRAP} scores at baseline and post-simulation on the Fear IRAP in Experiment 4

Figure 16: Mean D_{IRAP} scores at baseline and post-simulation on the Fear IRAP for high and low CAPE groups scores in Experiment 4

Figure 17: Schematic representation of the four trial-types presented in the Normality IRAP in Experiment 5

Figure 18: Mean D_{IRAP} scores for the two groups on the Normal IRAP in Experiment 5

Figure 19: Mean D_{IRAP} scores on the Normal IRAP for voice hearers and high and low CAPE groups scores in Experiment 5

Figure 20: Schematic representation of the four trial-types presented in the Self (left) and Others (right) IRAPs in Experiment 6

Figure 21: Mean D_{IRAP} scores for the two groups on the Self and Others IRAPs in Experiment 6

Figure 22: Mean D_{IRAP} scores for the two AAQ-II avoidance levels on the Self and Others IRAP in Experiment 6

Figure 23: Mean D_{IRAP} scores for the three groups on the Fear IRAP in Experiment 7

Figure 24: Mean D_{IRAP} scores on the Fear IRAP for voice hearers and high and low CAPE groups in Experiment 7

Figure 25: Mean D_{IRAP} scores for the two AAQ-II inflexibility levels on the Fear IRAP in Experiment 7

Figure 26: Mean D_{IRAP} scores for AHRS perceived controllability on the Fear IRAP in Experiment 7

Figure 27: Schematic representation of the four trial-types presented in the Valence (left) and the Acceptance (right) IRAPs presented in Experiment 8

Figure 28: Mean D_{IRAP} scores for the three groups on the Valence IRAP in Experiment 8

Figure 29: Mean D_{IRAP} scores for the three DASS distress groups on the Valence IRAP in Experiment 8

Figure 30: Mean D_{IRAP} scores for the two AAQ-II inflexibility levels on the Valence IRAP in Experiment 8

Figure 31: Mean D_{IRAP} scores for AHRS perceived controllability on the Valence IRAP in Experiment 8

Figure 32: Mean D_{IRAP} scores for the three groups on the Acceptance IRAP in Experiment 8

Figure 33: Mean D_{IRAP} scores for the three DASS distress groups on the Acceptance IRAP in Experiment 8

Figure 34: Mean D_{IRAP} scores for the two AAQ-II inflexibility levels on the Acceptance IRAP in Experiment 8

Figure 35: Mean D_{IRAP} scores for AHRS perceived controllability on the Acceptance IRAP in Experiment 8

List of Appendices

Appendix 1: The Community Attitudes to Mentally Ill Questionnaire

Appendix 2: Stigmatising Attitudes Believability Questionnaire

Appendix 3: Depression Anxiety and Stress Scales

Appendix 4: Participant Consent Form

Appendix 5: Illness Questionnaire

Appendix 6: Auditory Hallucinations Rating Scale

Appendix 7: Community Assessment of Psychic Experiences Questionnaire

Appendix 8: Acceptance and Action Questionnaire II

Appendix 9: Automatic Thoughts Questionnaire

Abstract

Abstract

The current programme of research had two main aims. First, in response to a gap in the literature on implicit measures in the context of psychosis, the current thesis sought to determine the utility of the Implicit Relational Assessment Procedure (IRAP) in this regard ($N=309$). We opted to focus specifically on the IRAP because of the level of concept precision it has demonstrated previously in the context of depression, cocaine dependence and OCD. Second, we used the IRAP as a step toward bringing a broad, functional approach to understanding psychosis, by focusing on the specific phenomenon of voice hearing. On this path, we created a taxonomy of the critical features of voice hearing as identified in the literature (i.e. stigma, locus of control, evaluations of voices, normality of voices, acceptance of self and other people hearing voices, fear of voices, valence and acceptance of voices). It was our hope that outcomes in this regard might lead us one step closer to a functional analytic understanding of the psychological nature of the experience of hearing voices.

The experimental designs and analytic strategies adopted in Experiments 1-3 were identical. All three studies involved control (non-voice hearing) participants completing one IRAP each. The six IRAPs employed across the studies had broadly similar trial-types that juxtaposed positive versus negative evaluations of various aspects of psychological suffering. Participants subsequently completed a battery of explicit measures which assessed psychological well-being and stigma.

Experiment 1 investigated the broad stigmatisation of psychological suffering using three IRAPs ($n=36$ undergraduate students), namely a Depression IRAP, an Anxiety IRAP and a Mental Illness IRAP. Participants were assigned into one of the

three IRAPs, in which they were required to relate suffering-as-positive and suffering-as-negative on alternative trial blocks.

Overall, the three IRAPs in Experiment 1 produced similar patterns of responding, which, surprisingly, indicated that all three forms of suffering (i.e. depression, anxiety and mental illness) were implicitly evaluated as both positive *and* negative. Similarly, all three IRAPs also showed significant pro-normality effects, although an anti-normality effect was recorded on the Mental Illness IRAP. The correlations revealed that IRAP trial-types predicted some aspects of explicit. For example, positivity toward depression correlated with low anxiety and low stigma, while negativity correlated with stigma. Positivity toward normality also correlated with high distress. And pro-normality effects in the context of anxiety correlated with overall distress. Overall, these correlations suggested that psychological distress impacts upon implicit evaluations to suffering and normality, and thus, may influence stigmatisation.

Experiment 2 investigated the locus of control in psychological suffering versus illness ($n=40$). We developed two IRAPs, namely an Illness IRAP and a Weakness IRAP, to investigate internal versus external locus of control. Participants were assigned into one of the two IRAPs, in which they were required to relate suffering-as-positive and suffering-as-negative on alternative trial blocks.

Both IRAPs produced similar patterns of responding, in which suffering (as illness or weakness) was evaluated as both positive *and* negative. Significantly strong positivity toward health was also recorded. The correlations revealed that positivity toward health correlated with high stigma on the Illness IRAP and high distress and stigma on the Weakness IRAP, on which negativity toward weakness also correlated with low distress.

Experiment 3 investigated stigmatisation of hearing voices by a non-voice hearing student population using a simple Evaluation IRAP ($n=29$). On the IRAP, participants were required to relate voices-as-positive and voices-as-negative on alternative trial blocks. Overall, hearing voices was positively evaluated. Surprisingly, implicit positivity toward voice hearing correlated with high stigma, but again negativity correlated with high distress.

Experiment 4 investigated potential changes in implicit stigmatisation or fear of hearing voices in non-voice hearing student participants subjected to a hearing voices simulation and a Fear IRAP ($N=28$). On the IRAP, participants were required to relate voices-as-positive and voices-as-negative on alternative trial blocks. Subsequently, participants were exposed to a hearing voices simulation procedure, prior to completing the IRAP for a second time. Participants subsequently completed a battery of explicit measures. At baseline and post-simulation, hearing voices was implicitly evaluated as both positive *and* fearful, although positivity toward voices reduced and fear increased after the simulation.

Experiment 5 investigated implicit evaluations of non-clinical voice hearers' and non-voice hearing control participants (all from a student population) using the Normality IRAP ($n=36$). On the IRAP, participants were required to relate voices-as-normal and voices-as-abnormal on alternative trial blocks. Participants subsequently completed a battery of explicit measures.

Both non-clinical voice hearers and controls implicitly evaluated voice hearing as normal, and most importantly this effect was stronger for the voice hearers. Evaluations of voice hearing as normal correlated with behavioural engagement with voices as well as voice acceptance, while evaluations of voice

hearing as abnormal correlated with high voice benevolence and high emotional engagement with voices and other psychotic-like experiences.

Experiment 6 investigated the potential stigma of non-voice hearing participants and non-clinical voice hearers (all from a student population) toward voices as heard by the self (i.e. “If I heard voices”) and others (i.e. “If Other People heard voices”), and presented these two groups with a Self IRAP and an Others IRAP ($n=48$). Participants were assigned into one of the two IRAPs, in which they were required to relate voices-as-positive and voices-as-negative on alternative trial blocks. Participants subsequently completed a battery of explicit measures.

Overall, hearing voices was evaluated positively by both groups on both IRAPs, although control participants were more positive on the Self IRAP. Interestingly, control participants were also more fearful on the Self IRAP, while the voice hearers were more fearful on the Others IRAP. Furthermore, voice hearers who were less positive on the Self IRAP correlated with high depressive psychotic-like symptoms and those who were fearful on the Others IRAP correlated with low voice acceptance.

Experiment 7 attempted to explore implicit evaluations by using a more broadly fear-based IRAP, and most notably by now including our first recruitment of clinical voice hearers ($n=37$). Clinical voice hearers were recruited from a psychiatric facility and the non-clinical voice hearers and non-voice hearing controls were recruited from a student population. On the IRAP, participants were required to relate voices-as-positive and voices-as-negative on alternative trial blocks. Participants subsequently completed a battery of explicit measures.

As expected, controls and non-clinical voice hearers showed implicit positivity and only marginal fear on the IRAP. But, most interestingly, the clinical

voice hearers were much less positive in this regard. The correlations also showed that implicit fear correlated with higher voice severity. Once again, these effects supported those we had observed previously, but critically showed that the clinical voice hearers were implicitly less positive than the other two groups.

Experiment 8 sought to further explore the potentially different reactions of clinical and non-clinical groups to voices, and in doing so we attempted to parse out emotional versus behavioural responses ($n=55$). In a Valence IRAP, participants were required to relate voices-as-positive and voices-as-negative on alternative trial blocks, and an Acceptance IRAP required to relate avoidance *and* acceptance of positive and negative voices on alternative trial blocks. Clinical voice hearers were recruited from a psychiatric facility and the non-clinical voice hearers and non-voice hearing controls were recruited from a student population. Participants subsequently completed a battery of explicit measures.

As expected, controls evaluated voice hearing positively on the Valence IRAP, but for the first time, non-clinical voice hearers showed implicit negativity, and we had now recorded for the second time, negativity by the clinical voice hearers. Clinical and non-clinical voice hearers implicitly accepted positive voices and avoided negative voices. Furthermore, acceptance of positive voices correlated with high psychological inflexibility, and acceptance of negative voices correlated with overall voice acceptance. This study provided the level of psychological precision that we had been working towards using the IRAP throughout the thesis.

The refinements and the systematic manipulations of the IRAP have led us steadily from simple assessments of valence by control participants to distinguishing between the different types of reactions that different groups of voice hearers might

have toward different types of voices, which largely coincides with the psychosis literature and thus suggests the validity of the IRAP within this domain.

Chapter 1

General introduction

General Introduction

Across almost all domains of experimental clinical psychology, the primary approach to exploring psychological suffering relies heavily on the direct measurement of symptoms (i.e. explicit measures and clinical interviews). This has been the dominant approach to clinical measurement and is supplemented only with (more indirect) physiological measures, such as galvanic skin response (GSR), heart rate and neurophysiological measures. But researchers, in the last decade, have increasingly emphasised the limitations of these direct, and even indirect, forms of measurement (De Houwer, 2006). Key to these concerns is the potentially problematic assumption that a direct measure, for example, is *actually* measuring the construct of interest. For instance, does a score on a depression scale actually reflect an *individual's* level of depressive suffering or does it act as a proxy measure of the *common* topographies that are often characterised as depressive behaviours (e.g. Barnes-Holmes, Hussey, McEnteggart, Barnes-Holmes, & Foody, in press)? Furthermore, explicit measures rely on an individual's *willingness* and *ability* to report private information accurately (Greenwald et al., 2002). Indeed, self-presentation biases and social desirability are likely to influence reporting, in terms of key information being withheld or distorted (Nisbett & Wilson, 1977; Paulhus, 2002; Wilson & Dunn, 2004). As a result of these factors, it is often difficult for basic and clinical researchers to isolate the phenomena of key concern and thus to identify the relevant *functional* processes at play in specific patterns of psychological suffering.

Using Implicit Measures to Study Clinical Phenomena

In an attempt to overcome these limitations, researchers have devoted increasing attention to the study and use of *implicit* measures to supplement data gathered by the more traditional explicit measures. For example, there are now over 40 published studies that have used the Implicit Association Test (IAT) to investigate clinically-relevant phenomena. In short, implicit procedures aim to measure target behaviour as it occurs, rather than asking participants to report on this behaviour (De Houwer, 2006). There are numerous other types of implicit measure that include: the affective priming task (e.g. Fazio, Jackson, Dunton, & Williams, 1995); the Extrinsic Affective Simon Task (EAST; e.g. De Houwer, 2003) and the word association task (e.g. Stacy, 1997), some of which, such as the EAST, have been used to study clinical phenomena. However, it is a reasonable assessment of the relevant literature to conclude that these various procedures have been significantly overshadowed by the wealth of research and impressive findings associated with the IAT.

In simple terms, the IAT asks participants to pair words or images, generally in an evaluative manner, under time pressure (e.g. insects-bad, flowers-good). The fundamental assumption of the IAT is that differences in reaction times (response latencies) of responding to the paired stimuli represent associations between these stimuli in memory. For example, Nock and colleagues (2010) assessed suicidal behaviours in 157 participants in an ER setting using a Death/Suicide IAT. This comprised of pairing words representing death or suicide (e.g. “suicide”) versus words representing life (e.g. “survive”) with self-related words (i.e. “myself”) and others-related words (e.g. “they”). The results indicated that participants who had attempted suicide responded significantly more quickly to the death-self pairings

than participants who had not attempted suicide but who were also clinically distressed. Even more interestingly, participants with these significantly shorter latencies on the death-self pairings were six times more likely to attempt suicide in the following six months. The IAT has similarly been used in the implicit assessment of other clinically relevant domains, such as anxiety -- again with impressive results, particularly in terms of clinically relevant predictive validity (Egloff & Schmukle, 2002; Teachman & Woody, 2004). In addition, it is now well established that the IAT has very sound psychometric properties (e.g. Greenwald, Poehlman, Uhlmann, & Banaji, 2009).

The Implicit Relational Assessment Procedure (IRAP). Although the IRAP has considerable structural overlap with the IAT, it hails from the behaviour analytic tradition, not normally concerned with group designs or procedures (Barnes-Holmes et al., 2006). However, the IRAP was developed specifically from Relational Frame Theory (RFT), a modern functional analytic approach to human language and cognition (Hayes, Barnes-Holmes, & Roche, 2001).

According to RFT, language and cognition emerge from our ability to relate stimuli in ways that are *not* based on their formal properties (e.g. shape or colour), but are based instead on arbitrarily applicable relations with other stimuli. For example, when offered one of two coins, a young child will likely select the larger, because of the superiority in physical size. However, if for example the coins were a 10-cent euro coin and a 5-cent coin, an older child will likely select the 10-cent coin, because even though it is physically smaller, the verbal community has established (arbitrarily) that it is greater in monetary value. Hence, the verbal functions of the smaller coin, as established by the social community, dominate the functions of physical superiority in terms of the child's responding. For RFT, it is this process

(referred to as arbitrarily applicable relational responding) which is key to verbal sophistication and is involved in *all* verbal behaviour, as established through our developmental histories (Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010).

For RFT, this history is established by exemplar training and natural language interactions, through which individuals learn to relate stimuli and events never before related together, but one can do so based indirectly from relevant histories of relating stimuli in similar ways. From a measurement perspective, RFT is interested in targeting these relational responses directly and exploring the types of verbal histories that give rise to these repertoires. And the IRAP was designed specifically for this purpose (Barnes-Holmes et al., 2010).

The IRAP shares many methodological features with the IAT. For example, it is an automated reaction-time based measure in which participants pair words and/or pictures. Its basic assumptions also match the IAT's in that participants should respond more quickly to pairings that are consistent with their pre-experimental verbal histories than pairings that are inconsistent. To illustrate, Barnes-Holmes et al. (2006) administered a simple IRAP comprising of the sample word stimuli "pleasant" and "unpleasant", pleasant-related target stimuli (e.g. "love" and "peace") and unpleasant-related target stimuli (e.g. "abuse" and "crash") and the relational terms "similar" and "opposite" as response options. On each trial, participants were presented with a sample, a target stimulus and the two relational response options. On blocks of trials deemed consistent, participants were required to respond with "similar" during pleasant-pleasant (e.g. pleasant-love-similar) and unpleasant-unpleasant (e.g. unpleasant-abuse-similar) trial-types and with "opposite" during pleasant-unpleasant and unpleasant-pleasant trial-types. On inconsistent blocks, participants were required to respond with "similar" for pleasant-unpleasant

and unpleasant-pleasant trial-types and “opposite” for pleasant-pleasant and unpleasant-unpleasant trial-types. The standardised difference score between response latencies on consistent and inconsistent blocks of trials generates four D_{IRAP} scores for each trial-type (i.e. pleasant-pleasant, pleasant-unpleasant, unpleasant-unpleasant and unpleasant-pleasant). In the original 2006 study, Barnes-Holmes et al. found, as expected, larger D_{IRAP} scores for trials that were consistent with participants’ pre-experimental verbal histories (e.g. pleasant-pleasant and unpleasant-unpleasant) than those that were inconsistent. In numerous studies subsequently, the IRAP has also demonstrated good reliability and predictive validity (Carpenter, Martinez, Vadhan, Barnes-Holmes, & Nunes, 2013; Fischer, 2013).

Using the IRAP to study clinical phenomena. As the body of supporting evidence for use of the IRAP grows steadily (there are now over 50 published empirical articles), the measure has come to be used increasingly in the study of clinical phenomena. For example, Hussey and Barnes-Holmes (2012) assessed depressive behaviours in 76 non-clinical participants using a Depression IRAP. This comprised of pairing short statements about success (e.g. “When things go well”) or failure (e.g. “When things go badly”) with statements about *reactions* to success or failure as positive (e.g. “I feel happy”) or negative (e.g. “I feel sad”). The results indicated that participants who were ‘above normal’ on a standardised clinical explicit measure of depression showed stronger responding to success-as-positive (i.e. implicitly evaluating success positively) following a mood induction procedure than participants who were in the normal range.

One significant advantage of the IRAP over the IAT is its ability to provide greater precision in terms of understanding the observed patterns of responding. That

is, while the IAT identifies *associations* between pairings, the IRAP also *specifies the nature of* these pairings. Consider the study by Nicholson and Barnes-Holmes (2012b) with 33 non-clinical participants who completed two IRAPs, both of which assessed disgust toward pleasant pictures (e.g. neatly folded towels) or unpleasant pictures (e.g. a dirty toilet). Specifically, one IRAP assessed disgust *propensity* (i.e. the tendency to experience disgust), while the other assessed disgust *sensitivity* (i.e. how negatively a disgust experience is appraised). In simple terms, the disgust propensity IRAP measured *emotional* reactions, while the sensitivity IRAP measured *behavioural* reactions. Participants also undertook a series of behavioural approach tasks (BAT) and explicit measures. The results demonstrated that while responding on both IRAPs predicted obsessive compulsive tendencies on explicit measures of obsessive compulsive disorder (OCD), *only the sensitivity* IRAP predicted avoidance behaviours on the BAT. That is to say that the behavioural reaction to the event predicted actual behaviour, while the emotional reaction to it did not. The IRAP's ability to separate these two constructs of the same overarching feature (disgust) highlighted the potential promise of the measure in clinical domains. Furthermore, the IRAP has recently been shown to predict treatment outcomes with cocaine dependence (Carpenter et al., 2013).

One of the key domains of psychological suffering that lays claim to perhaps the largest published literature, but which has attracted little or no interest from researchers of implicit measures, is *psychosis*. While this lack of interest may simply reflect the perceived complexity of the patterns and behaviours that characterise psychotic-like suffering, it nonetheless renders this area highly dependent on explicit measures to be completed by individuals who are often deemed to have reduced capacity to respond coherently and who are characterised as lacking insight into their

situation (see Kim et al., 2010). If indeed these latter views are correct, then psychotic-like suffering would be a key area in which implicit measures may play an important role.

At the time of writing, the current author found only a *handful* of published studies that used an implicit measure in the context of psychosis. For example, McCulloch, Clare, Howard, and Peters (2006) used the Stroop task to investigate implicit depression and delusions in late-onset psychosis, and found implicit biases toward depression-related and age-related words. More recently, Wiffen and colleagues (2014) used the Stroop task to explore the relationship between implicit biases and insight into one's suffering using psychosis-related stimuli (e.g. "crazy"), general negative stimuli (e.g. "cancer") and neutral stimuli (e.g. "oyster"). The authors reported that psychosis-related words were less threatening and less self-relevant in psychosis sufferers with *less* insight into their own situation. Other researchers have used the self-referent incidental recall task to investigate whether a depressive-type self-concept accompanied delusions, but no implicit negative schema were observed (Vázquez, Díez-Algeria, Hernández-Lloreda, & Moreno, 2008). Finally, Soler, Ruiz, Vargas, Dasí, and Fuentes (2011) used the word fragmented completion task to explore implicit memory in schizophrenia, but no specific deficits were recorded. A number of additional IAT studies pertaining to the stigmatisation of psychosis are summarised later in the chapter.

Understanding "Psychosis"

While an extensive review of the vast psychosis literature is well beyond the scope of the current thesis (but see Boyle, 2004), it is useful to provide a brief

summary of the core features of this type of psychological suffering, as highlighted in the literature.

According to DSM-V, psychotic disorders are typically defined by the presence of “positive” psychotic-like symptoms (i.e. hallucinations and/or paranoid delusions). Under the umbrella of psychotic disorders, the most common diagnosis given to sufferers with these symptoms is schizophrenia. Within DSM-V, a diagnosis of schizophrenia *requires the following five criteria*: 1. A significant presence of two or more of the following characteristics during a one-month period: delusions¹; hallucinations¹; disorganised speech; disorganised behaviour or catatonic behaviour; or negative symptoms (i.e. diminished emotional expression or avolition); 2. Social and/or occupational dysfunction; 3. A disturbance which persists for at least six months, including one month of the symptoms outlined in criterion 1; 4. No presence of other disorders which may explain these symptoms (i.e. schizoaffective, depressive or bipolar disorder); and 5. The disturbance is not attributable to other physiological effects (i.e. drug use).

In spite of these precise requirements, the reliability and utility of the DSM criteria for schizophrenia have been severely criticised (Kirk & Kutchins, 1994; Van Os et al., 1999). As an alternative, many authors have proposed that psychotic-like experiences (i.e. hallucinations and delusional beliefs) lie on a *continuum with normal experiences*. Indeed, this latter view is supported by a wealth of literature indicating that these experiences also present in other patterns of psychological suffering *and* in non-clinical populations (Johns & Van Os, 2001). Given the breadth of the label of psychosis, and even schizophrenia, and the potential benefits of exploring key features that may not be specific to these patterns of suffering,

¹ At least one of these two symptoms must be present for a diagnosis of schizophrenia

researchers have begun to investigate specific psychotic-like experiences. And among these, the experience of hearing voices has attracted strong attention. This impetus is likely due to two related facts. 1. It is now established that voices are a very commonly reported symptom in all clinical contexts, not only in diagnoses of psychosis but also in other psychiatric diagnoses. 2. Social movements, such as the Hearing Voices Movement, have grown rapidly and are now powerful advocates for social change (Bentall, 2004; Corstens, Longden, McCarthy-Jones, Waddingham, & Thomas, 2014). The focus of the current thesis is on the phenomenon of hearing voices.

Voice hearing research. It is perhaps not surprising that hearing voices is highly prevalent (approx. 70%) in individuals with a diagnosis of schizophrenia (Sartorius et al., 1986). On balance however, it is also prevalent in other psychiatric diagnoses, such as: borderline personality disorders (approx. 32%; e.g. Slotema et al., 2012); dissociative disorders (approx. 70-90%; e.g. Dorahy et al., 2009); post-traumatic stress disorder (PTSD; approx. 50%), and bipolar disorders (approx. 7%; e.g. Blakemore, Smith, Steel, Johnstone, & Frith, 2000). And critically, hearing voices is *also* common (approx. 10-15%) in individuals with *no* clinical diagnosis, social and/or occupational dysfunction or psychological distress (Beavan, Read, & Cartwright, 2011; Eaton, Romanoski, Anthony, & Nestadt, 1991; Rössler et al., 2007; Sommer et al., 2010; Tien, 1991).

Consistent with the psychosis literature generally, *all* studies of voice hearing have relied *solely* on explicit measures (for both clinical and research purposes, see Kim et al., 2010; Ratcliff, Farhall, & Shawyer, 2010). Explicit measures of voice hearing primarily focus on the phenomenological features of voices, appraisals about voices or reactions to voices. For example, consider the Beliefs about Voices

Questionnaire (BAVQ) which assesses voice appraisals, as well as emotional and behavioural responses to voices (Chadwick & Birchwood, 1995). Also consider the Auditory Hallucinations Rating Scale (AHRs: Haddock, McCarron, Tarrier, & Farragher, 1999) which assesses various dimensions of voices (i.e. frequency, duration, locations, loudness, disruption, amount and intensity of distress, beliefs about origins, amount and degree of negative content and controllability).

While these and similar studies may be perceived as limited because they relied entirely on explicit measures, it is important to recognise that this type of research has played an important role in understanding and assessing the key features of the voice hearing experience in both clinical and non-clinical populations. Consider the following evidence. It is well established that voices are better understood in terms of an individual's relationship with their voices (Birchwood & Chadwick, 1997; Chadwick & Birchwood, 1994). And, appraisals of voices directly influence an individual's affect and behaviour, independently of the presence of voice hearing and voice content (Birchwood & Chadwick, 1997; Peters, Williams, Cooke, & Kuipers, 2012; van der Gaag, Hageman, & Birchwood, 2003). Depression, in particular, is frequently associated with disability and morbidity in distressed voice hearers (Birchwood & Chadwick, 1997; Birchwood, Iqbal, & Upthegrove, 2005). And, there is increasing evidence that experiential avoidance is associated with: delusional ideation; distress; disability; paranoia; poorer self-esteem; and depression and anxiety following psychosis (Goldstone, Farhall, & Ong, 2011; Morris, Garety, & Peters, 2014; Udachina, Varese, Myin-Germeys, & Bentall, 2012; White et al., 2011).

In spite of a substantive body of research using the IAT in other clinical domains and a growing similar body of work on the IRAP, there appear to be only a

handful of published IAT studies and no IRAP studies in the context of psychotic-like experiences. In fact, there are no studies published that have used implicit measures to investigate the experience of voice hearing. And although almost all schools of thought in psychology have offered comprehensive, eloquent and often overlapping accounts of psychotic experiences, including voice hearing, very little has emerged from the functional-analytic community. For us, this approach would potentially offer an understanding of why and how voice hearing occurs (i.e. what are the key variables and processes involved) and is maintained (i.e. the psychological functions served by these behaviours). In our review of the literature on psychosis, we identified only one such account, by Rosenfarb (2013) which adopted a very traditional behavioural view in which voices were believed to emerge when “*other, more potent and appropriate reinforcers are unavailable*” (p. 933). Accordingly, this loss of reinforcement forces the individual to redirect his/her focus inwards, which in turn minimises the impact of other aversive experiences and may itself be reinforcing.

In response to this gap in the literature, the current thesis sought to determine the utility of the IRAP as a methodological step towards a broad, functional approach to voice hearing, using a taxonomy of the critical features of voice hearing as identified in the literature (i.e. stigma, locus of control, evaluations of voices, normality of voices, acceptance of self and other people hearing voices, fear of voices, valence and acceptance of voices). We opted to focus specifically on the IRAP because of the level of concept precision it had demonstrated previously in the context of depression, cocaine dependence and OCD. It was our hope that outcomes in this regard might lead us one step closer to a functional analytic understanding of the psychological nature of the experience of hearing voices. Outlines of each of the

critical features of voice hearing as identified in the literature are provided in the sections below.

Stigmatisation of psychological suffering. Attitudes toward psychological suffering, from others and from sufferers themselves, have generated much research interest in recent years (e.g. Bentall, 2003; Link, Yang, Phelan, & Collins, 2004; Schulze, 2007). For example, Link, Cullen, Struening, Shrout, and Dohrenwend (1989) proposed a labelling theory through which they argued that labels themselves are likely causes of stigma and discriminative behaviour. This is consistent with Szasz's long-standing (1960) view that discrimination towards others reflects one's own fear of belonging to an 'abnormal' group and being rejected by society. Indeed, researchers have found that the *specific* labels used to describe patterns of human suffering differentially influence the attitudes of non-sufferers (Crisp, Gelder, Rix, Mertzner, & Rowlands, 2000; Lauber, Nordt, Falcató, & Rössler, 2003; Reavley & Jorm, 2011; Stier & Hinshaw, 2007).

Only a handful of studies have used the IAT to investigate *implicit* stigma of psychological suffering. For example, Peris, Teachman, and Nosek (2008) examined whether implicit stigma in those with differential levels of mental health training influenced their clinically-relevant decision-making. And the results indicated that individuals with the most training in mental health portrayed the least stigma, but this bias nonetheless predicted over-diagnosis. In a study of implicit stigma against depression and schizophrenia as measured in a pre- and post-stigma intervention, attitudes held by both medical and psychology students were more negative for schizophrenia than depression (Lincoln, Arens, Berger, & Rief, 2008).

Researchers have also suggested that stigmatised individuals view themselves as inferior and different to the general population (Link & Phelan, 2001).

In a study of implicit self-stigmatising behaviour, Teachman, Wilson, and Komarovskaya (2006) investigated stigma against psychological suffering versus physical illness in individuals with and without a psychiatric diagnosis, and found negative implicit biases and beliefs about the blame and helplessness for psychological suffering, but not for physical illness. Rüsçh, Corrigan, Todd, and Bodenhausen (2010) also found that implicit self-stigma in individuals with a clinical diagnosis predicted a lower quality of life, independently of presence of depression, the type of diagnosis and demographics. On balance, less self-stigmatising behaviour has been associated with *higher* social functioning (Mersh, Jones, & Oliver, in press).

Locus of control. The concept of locus of control may have potential implications for the aetiology of stigmatising attitudes toward psychological suffering. For example, Rotter (1966) proposed that individuals who believe that reinforcement is directly related to their own behaviour are more likely to have an *internal* locus of control, whereas individuals who believe that outcomes are mediated by external factors and others are more likely to have an *external* locus of control. Levenson (1973) added that locus of control varies along three independent dimensions of internalisation, chance and powerful others.

In a study of the relationship between locus of control and stigmatisation towards others with psychological suffering, Benkman (1971) reported that mental health volunteers with a greater *internal* locus of control endorsed more positive attitudes toward mental illness than those with an *external* locus of control. Indeed, Morrison, D'man, and Drumheller (1994) reported that individuals with external locus of control had greater negative attitudes toward mental illness, whereas those with internal locus of control had greater positive attitudes. Similarly, Ryu (2001)

reported that internal locus of control significantly predicts positive attitudes toward mental illness.

Evaluation of voices. Studies of stigma often focus on the broad constructs of ‘mental illness’ or ‘schizophrenia’, rather than specific features or patterns of suffering. Hence, there is little empirical research on evaluations of specific phenomena. In cross-cultural research, Al-Issa (1977, 1995) investigated evaluations of hearing voices and reported large variations across cultures. Specifically, cultures that supported a flexible distinction between reality and fantasy typically had more positive evaluations of voice hearing experiences and individuals are actually *encouraged* to observe and increase their awareness of hallucinatory experiences and other private events. As a result, one would predict that voice hearers would experience less stigmatisation in these cultures. This contrasts sharply with cultures (most of which are Western) that advocate a *rigid* distinction between reality and fantasy, and thus evaluate voice hearing more negatively. In these contexts, detachment from reality is believed to be indicative of mental illness and thus voice hearers are more likely to be stigmatised.

A growing number of studies have examined stigma toward voices in the context of interventions that attempt to target the negative attitudes of mental health professionals. For example, voice hearing simulation procedures successfully reduce stigma and improve attitudes toward voices (Dearing & Steadman, 2009; Deegan, 1996; Wieland, Levine, & Smith, 2015; Wilson et al., 2009). Similarly, Sideras, Mckenzie, Noone, Dieckmann, and Allen (2015) reported decreases in negative attitudes, fear and behavioural intentions post-simulation, but this effect was only found with professionals with *less* rather than more clinical experience. Kidd, Tusaie, Morgan, Preebe, and Garrett (2015) also found that a voice hearing

simulation increased patience, tolerance, understanding and empathy toward voice hearers. However, this study also showed that post-simulation participants were *less* likely to acknowledge aspects of voices as indicators of recovery. Indeed, some studies have reported negative post-simulation outcomes, in which participants show increased social distance from voice hearers (Brown, Evans, Espenschade, & O'Connor, 2010; Kalyanaraman, Penn, Ivory, & Judge, 2010).

Appraisals of voices. Various authors have argued that voice appraisals play an important role in the distress associated with voice hearing (Johns et al., 2014; Mawson, Cohen, & Berry, 2010; Morrison, 2001; Romme & Escher, 1989). Voice appraisals simply refer to whether voice features are evaluated as positive or negative qualities and these appraisals vary considerably among voice hearers (Haddock et al., 1999; Woods, Jones, Alderson-Day, Callard, & Fernyhough, 2015). Several studies have also shown that appraisals of voices influence: level of perceived control; emotional reactions to the voices; voice frequency; and voice intensity (Al-Issa, 1995; Morrison, Nothard, Bowe, & Wells, 2004).

A key distinction that has emerged in the rapidly-growing research area on voice appraisals juxtaposes perceptions of benevolence, malevolence and omnipotence. In short, appraising voices as *benevolent* has been associated with voice engagement, appraising voices as *omnipotent* (or even fearful) has been associated with distress and appraising voices as *malevolent* has been associated with voice resistance (Chadwick & Birchwood, 1994; Gauntlett-Gilbert & Kuipers, 2005; Jackson, Hayward, & Cooke, 2011; Mawson, Berry, Murray, & Hayward, 2011; O'Brien & Johns, 2013; Peters et al., 2012). And these findings are independent of frequency of voices, severity of voices and intensity of voices (Peters et al., 2012).

Responding to voices: Clinical and non-clinical voice hearers. Various studies have reported differences in phenomenological voice features and responses to voices, among clinical and non-clinical voice hearers (for a review, see Johns et al., 2014). Overall, the voices of clinical voice hearers are often more negatively valenced, distressing, frequent, and intrusive than those heard by non-clinical voice hearers (Choong, Hunter, & Woodruff, 2007; Daalman et al., 2011; Diederens, Van Lutterveld, & Sommer, 2012; Hill, Varese, Jackson, & Linden, 2012). And indeed, it has been suggested that these differences likely contribute to the differential levels of distress and patterns of responding toward voices in clinical versus non-clinical voice hearers (Beavan & Read, 2010; Jenner, Rutten, Beuckens, Boonstra, & Sytema, 2008; Krabbendam, Myin-Germeys, Bak, & Van Os, 2005; Vaughan & Fowler, 2004).

Several studies have also shown that clinical and non-clinical voice hearers can be differentiated in terms of *emotional and behavioural* responding toward voices. For example, Brett et al. (2007) reported that non-clinical voice hearers have full voice acceptance and use less voice avoidance strategies, relative to clinical voice hearers. Furthermore, clinical voice hearers report less perceived controllability of voices and more maladaptive coping than non-clinical voice hearers (Andrew, Gray, & Snowden, 2008; Honig et al., 1998; Ward et al., 2014). Indeed, Daalman and colleagues (2011) also reported that less perceived control in turn predicted a clinical diagnosis of psychosis and increased contact with mental health services. Taken together, these findings may account for the higher levels of distress and ‘need for care’ among clinical voice hearers, relative to their non-clinical counterparts (Johns et al., 2014).

Rationale of Experiments 1-8

Having reviewed the literature on implicit measures, psychosis and voice hearing specifically, the current thesis set about investigating the implicit emotional and behavioural processes that may influence the experience of voice hearing. The primary aim was to determine the potential utility of the IRAP in this domain, but we commenced this endeavour more broadly by using the IRAP to explore stigma toward psychological suffering. We thereafter focused specifically on the key features of voice hearing as identified by the literature and systematically sought to use the IRAP to explore these phenomena. At one level, we simply hoped to contribute to the small literature on complementing traditional explicit measures with implicit measures in the study of psychotic-like experiences. But, at a broader level, we hoped that we would be able to take the first few small steps toward gaining a functional-analytic understanding of these psychological events, an approach that has remained almost completely untapped in the vast literature on the phenomenon known as ‘psychosis’.

Notably, the core features investigated here were not guided by the literature alone, but also through the systematic experimental investigations of the current thesis. That is, once each experiment was completed, the next functional analytic question was posed. This process progressed from implicit attitudes toward psychological suffering more broadly, toward more specific features of suffering, such as hearing voices. Once these baselines were observed, implicit self-stigma of clinical and non-clinical populations was investigated, which therefore facilitated more specific investigations about voices hearing, such as differences between clinical and non-clinical responding toward voices.

There are strong overlaps in the experimental designs and analytic strategies employed across the eight studies, which permit useful comparisons of the response patterns of the various groups and of the range of psychological phenomena investigated with the implicit and explicit measures. The experimental design and analytic strategy adopted in Experiments 1-3 are identical. That is, all three studies involve control (non-voice hearing) participants completing only one IRAP each and the six IRAPs employed across these three studies have largely similar trial-types that juxtapose positive versus negative evaluations of various aspects of psychological suffering (i.e. depression, anxiety, mental illness, illness, weak-mindedness and voice hearing).

Analyses of variance are then used to assess the extent to which trial-type performances vary across these IRAPs. Experiments 1 and 2 also involve distress level analyses on the Depression Anxiety and Stress Scales (DASS; Lovibond & Lovibond, 1995) data obtained with the control participants to investigate the potential influence of distress on the IRAP effects. In a similar vein, Experiment 3 also involves distress level analyses of the data from control participants. However, because the IRAP here was the first to explore voice hearing, it was more appropriate to present the Community Assessment of Psychic Experiences (CAPE; Stefanis et al., 2002) than the DASS, hence the distress level analyses in Experiment 3 are conducted on the CAPE data. Experiments 1-3 also employ correlational analyses to investigate the potential predictive validity of the IRAPs used here.

Experiment 4 differs in several ways from Experiments 1-3. Although it involves control participants presented with a single IRAP that largely resembled the voice hearing IRAP from Experiment 3, this study employs a pre-post experimental design to determine the potential impact of a voice hearing simulation procedure on

participants' fearful responding on the IRAP. Analyses of variance are then used to assess the potential impact of the simulation on the IRAP and distress level analyses are conducted on the CAPE data. Experiment 4 again employs correlational analyses to investigate the potential predictive validity of the IRAP.

Experiments 5 and 6 were the first in the thesis to involve between groups analyses because both studies compare control participants with non-clinical voice hearers (i.e. on the normality of voices and evaluations of the self and others hearing voices). However, these two studies vary in several ways. The distress level analyses in Experiment 5 involve the CAPE, while those in Experiment 6 involve the Acceptance and Action Questionnaire-II (AAQ-II; Bond et al., 2011), but both again are employed to determine their potential influence on IRAP effects. Experiments 5-6 again employ correlational analyses to investigate the potential predictive validity of the IRAPs.

Experiments 7 and 8 also involve between groups analyses but compare *three* groups: control participants, non-clinical voice hearers and clinical voice hearers (i.e. on fear and valence and acceptance of voices). Again, distress level analyses are conducted, but the measures varied according to the content of the IRAPs. In Experiment 7, the distress level analyses involve data from the CAPE, the AAQ-II and the AHRS, while Experiment 8 analyses distress in terms of the DASS, the AAQ-II and the AHRS. Once again, these analyses seek to determine the potential influence of distress on IRAP effects. Once again, Experiments 7-8 employ correlational analyses to investigate the potential predictive validity of the IRAPs.

Chapter 2

Experiments 1 and 2

Assessing stigmatisation and locus
of control in psychological suffering
with a non-clinical sample

Experiment 1

Assessing Implicit Stigmatisation to Psychological Suffering

Experiment 1 sought to determine the utility of the IRAP in the broad study of stigma toward different patterns of psychological suffering. Specifically, the concepts of depression (using the Depression IRAP), anxiety (using the Anxiety IRAP) and mental illness more broadly (using the Mental Illness IRAP) were targeted, and each of these three was juxtaposed in the IRAPs against normality. These three broad concepts were chosen because they are used with high frequency in everyday dialogue about psychological suffering and are typically subjected to negative evaluations. In this simple and first exploratory IRAP, we sought to determine whether negative attitudes or stigma toward these three concepts relative to normality would be observed as IRAP effects. In short, would participants more readily negatively evaluate depression, anxiety, or mental illness, than normality? We hypothesised that this precise pattern would be the case.

Method

Setting

The current study was conducted in an experimental cubicle at the Department of Psychology, Maynooth University (MU). All participation was on an individual basis. On average, experimental sessions lasted between 30 and 60 minutes, and all participation was completed in one session. The experimenter interacted with participants only during instructional phases of the IRAP and remained seated behind participants at all other times.

Participants

The current study involved a group of undergraduate students recruited from MU. Sixty-three individuals participated, 42 were female and 21 were male, with an age range of 20 to 32 years and a mean age of 26.33 years and a standard deviation of 2.31 years.

Materials

Explicit measures. Two broad categories of explicit measures were administered. The first series of measures assessed stigma (the CAMI and the SAB, see below). The second measure more broadly assessed general psychological well-being (the DASS, see below).

The Community Attitudes to Mentally Ill (CAMI; Taylor & Dear, 1981).

The CAMI is a 40-item measure of attitudes towards individuals with mental health difficulties distributed across four subscales: authoritarianism, benevolence, social restrictiveness and community mental health ideology (e.g. “as soon as a person shows signs of mental disturbance, they should be hospitalised”). All items are rated on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). The CAMI yields a total score for each subscale with a maximum of 50 and a minimum of 10. Higher scores on the authoritarian and social restrictiveness subscales indicate more stigmatising attitudes, while higher scores on the benevolence and mental health ideology subscales indicate more accepting attitudes. A total CAMI score is calculated by summing the scores on authoritarian attitudes and social restrictiveness, as well as summing the scores on benevolence and ideology, and then subtracting the latter from the former (i.e. $[A+R] - [B+I]$). Total CAMI scores range from -80 to 80, with a higher score indicating a higher stigmatising attitude. This

scale has demonstrated adequate internal consistency on each subscale: authoritarianism ($\alpha=.68$); benevolence ($\alpha=.76$); social restrictiveness ($\alpha=.80$); and community mental health ideology ($\alpha=.88$; Taylor, Dear, & Hall, 1979). See Appendix 1.

Stigmatising Attitudes Believability (SAB; Masuda, Price, Anderson, Schmertz, & Calamaras, 2009). The SAB is an 8-item measure of believability of negative statements about individuals with mental health difficulties (e.g. “those with psychological disorders are dangerous to others”). All items are rated on a 7-point scale from 1 (not at all believable) to 7 (completely believable). The SAB yields an overall score with a maximum of 56 indicating a high stigmatising attitude and a minimum of 8 indicating a low stigmatising attitude. This scale has demonstrated adequate internal consistency with an alpha coefficient of 0.78 (Masuda et al., 2009). See Appendix 2.

Depression Anxiety and Stress Scales (DASS-21; Lovibond & Lovibond, 1995). This 21-item DASS comprises three subscales that measure depression, anxiety and stress (e.g. “I found it hard to wind down”). All items are rated on a 4-point scale from 0 (did not apply to me at all) to 3 (applied to me very much or most of the time). Subscales are scored independently and indicate high or low levels of each type of distress (e.g. maximum depression score = 21 and minimum depression score = 0). This scale has demonstrated excellent internal consistency with an alpha coefficient of 0.93 for the total DASS score and the three sub-scales: depression ($\alpha =0.82$); anxiety ($\alpha =0.90$); and stress ($\alpha =0.93$, Henry & Crawford, 2005). See Appendix 3.

The Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2006). The IRAP was administered on a standard desktop computer. This was

used to present the instructions and stimuli, and to record responses. The current study involved three separate IRAPs (referred to as the Depression IRAP, the Anxiety IRAP and the Mental Illness IRAP) which assessed positive or negative evaluations of psychological suffering. The design of the IRAP is repeated measures as there are multiple presentations of the four trial-types.

The IRAPs contrasted psychological suffering with normality using, for example, the labels DEPRESSED PERSON IS and NORMAL PERSON IS (please note that capitalised words indicate actual stimuli used in the IRAP). Each trial-type presented one of these two types of category labels. These were accompanied by one of six positive (e.g. SANE) or six negative target stimuli (e.g. TROUBLED). The screen also presented two response options, TRUE and FALSE. Before each block of trials, the screen presented one of two rules for responding (e.g. PLEASE ANSWER AS IF DEPRESSED PEOPLE ARE NEGATIVE AND NORMAL PEOPLE ARE POSITIVE or PLEASE ANSWER AS IF NORMAL PEOPLE ARE NEGATIVE AND DEPRESSED PEOPLE ARE POSITIVE). A full list of label stimuli, target stimuli and response options for the IRAPs is provided in Table 1.

Table 1

Stimuli and Response Options of the three IRAPs employed in Experiment 1

Depression IRAP		Anxiety IRAP		Mental Illness IRAP	
Depressed Person is	Normal Person is	Anxious Person is	Normal Person is	Mentally Ill Person is	Normal Person is
Troubled	Composed	Panicky	Relaxed	Strange	Friendly
Inefficient	Efficient	Phobic	Rational	Crazy	Sane
Sad	Happy	Nervous	Calm	Sad	Happy
Unfocused	Focused	Worried	Mellow	Paranoid	Relaxed
Lazy	Energetic	Obsessive	Easy Going	Nervous	Calm
Mad	Sane	Uneasy	Composed	Phobic	Rational
True	False	True	False	True	False

Procedure

The current study comprised of three IRAPs: The Depression IRAP; The Anxiety IRAP; and The Mental Illness IRAP. For all participants, there were two stages, one involving the implicit measure (an IRAP) and the second involving the explicit measures. It is important to emphasise, therefore, that each participant only completed *one IRAP*, the Depression IRAP, the Anxiety IRAP *or* the Mental Illness IRAP, hence, approximately one third of the total group of participants completed each IRAP. Thus, the design of the study was between groups repeated measures. Participants were quasi-randomly assigned to one of the three IRAPs and all participants completed their respective IRAP prior to the explicit measures.

The verbal and automated instructions provided to participants for completing each IRAP were consistent with those in the most recently published IRAP research (e.g. Nicholson & Barnes-Holmes, 2012a). In short, these highlighted three key features of the task: the criterion for high levels of accurate responding (i.e. 80%), the criterion for responding very quickly (<2,000 ms.), and the fact that the ‘correct’ and ‘incorrect’ patterns of responding (depending upon the rule provided) alternated across blocks of trials. Similarly, the presentation features of the IRAP were identical to most recently published work in terms of: a maximum of four pairs of practice blocks (depending upon performance); three pairs of test blocks; 24 trials in every block; four trial-types; and two specified rules for responding. For all participants, blocks alternated between two patterns of responding according to the specified rules (e.g. responding as if depression is negative vs. responding as if depression is positive, see Figure 1). Blocks were counterbalanced across participants in terms of which rule was presented first (e.g. Rule A in the first block, Rule B in the second block, Rule A in the third block and so on). For illustrative

purposes, see Figure 1 for a schematic representation of the screen presentation of the Depression IRAP. The same format also applied to the Anxiety IRAP and the Mental Illness IRAP.

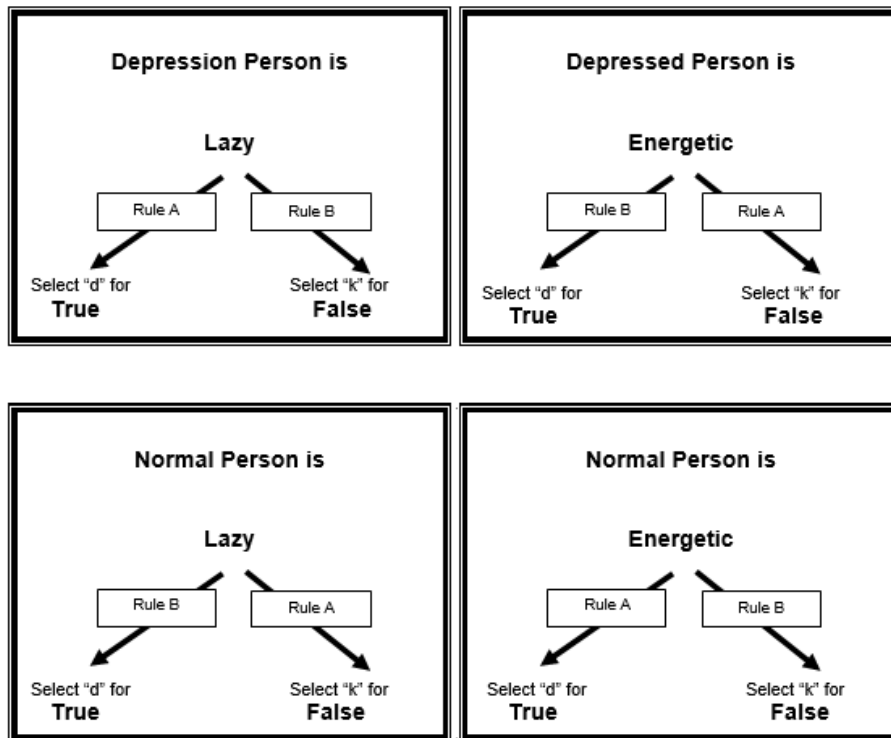


Figure 1. Schematic representation of the four trial-types presented in the Depression IRAP in Experiment 1. The arrows and text boxes did not appear on-screen, they indicate the correct responses for Rule A and Rule B blocks of trials. **The labels used for the four trial-types are as follows: Disorder-Negative (top-left), Disorder-Positive (top-right), Normality-Negative (bottom-left) and Normality-Positive (bottom-right).**

Once participants finished the IRAP, they completed the three explicit measures in a pre-determined sequence (CAMI, SAB and DASS).

Ethical Considerations

All aspects of Experiment 1 adhered to the ethical guidelines outlined by the British Psychological Society (BPS, 2009) and the Psychological Society of Ireland

(PSI, 2011), and received prior ethical approval from the Maynooth University Ethics Committee or underwent ethical review at the departmental level (Approval date: October 2011). The key steps taken to circumvent any potential ethical issues pertaining to the current study may be summarised as follows: 1. Each participant completed a consent form, which highlighted specific ethically-relevant features, provided details on the nature and aims of the research, and outlined freedom to withdraw at any point (see Appendix 4). 2. All data was anonymised and analysed at group level. 3. No participant who reported a history of clinical psychological distress in the explicit measures was exposed to any of the implicit measures. 4. The lack of distress or harm associated with the IRAP has been demonstrated by empirical evidence (Hussey et al., under submission). 5. All participants were fully debriefed and provided with researcher contact details. It is important to emphasise that no participant expressed any signs of distress prior to, during, or after involvement in the study.

Results

Analytic Strategy

To assess the potential effects of IRAP type on trial-type, a mixed between within ANOVA was conducted. Exploratory distress level analyses using the DASS were then conducted to investigate the potential influence of distress on level of stigmatisation as measured on each IRAP (i.e. whether there are stigmatising or non-stigmatising effects). This involved three mixed between within ANOVAs. Correlational analyses were then carried out to investigate the potential predictive validity of the IRAP in this context.

Before conducting the IRAP analyses, summaries of the explicit data are provided in the section below.

Explicit Measures Data

The means and standard deviations were calculated from each participant's responses on each of the explicit measures and data are summarised in Table 2.

Table 2

Explicit Data Summary for Experiment 1

Scales	Anxiety IRAP	Depression IRAP	Mental Illness IRAP	ANOVA
	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>F</i>
DASS				
DASS Total	45.54 (36.94)	31.7 (22.94)	54.76 (37.36)	.099
Depression	17.84 (18.44)	6.16 (5.56)	17.54 (15.96)	.072
Anxiety	7.1 (8.08)	6.76 (7.76)	11.38 (10.30)	1.361
Stress	20.00 (16.08)	18.76 (14.08)	25.84 (17.02)	.373
CAMI				
Authoritarian	18.08 (4.33)	17.15 (4.69)	18.31 (3.35)	.280
Benevolence	41.62 (3.25)	44.62 (4.61)	41.46 (5.16)	2.110
Social Restrictiveness	17.54 (2.67)	16.69 (4.96)	18.23 (3.24)	.549
Mental health Ideology	41.23 (3.32)	42.69 (5.68)	38.85 (5.57)	1.980
SAB Total	19.08 (7.59)	18.46 (6.06)	22.15 (8.12)	1.029

*Note. Maximum scores are: DASS total = 126; DASS subscales = 42; CAMI = 10-50; SAB = 56. Significance indicated by *($p < 0.05$).

For the DASS, means revealed that participants overall had low rates of depression, anxiety and stress. For the CAMI, means revealed that participants overall had low stigmatisation and high acceptance of mental illness. And for the SAB, means revealed low stigmatising attitudes toward mental illness. One-way

ANOVAs were conducted to investigate potential differences between the three groups of participants, but none were found (all p 's > 0.05).

IRAP Data

Scoring of the IRAP was always conducted using the standardised approach for transforming latency data into D_{IRAP} scores (see Nicholson & Barnes-Holmes, 2012a). That is, response latencies were the primary datum. These were recorded in milliseconds, commencing from the beginning of each trial-type to the point at which a participant emits a correct response. Response latencies were transformed into D_{IRAP} scores based on the D_{IRAP} algorithm (derived from a similar IAT-based algorithm by Greenwald, Nosek, & Banaji, 2003). This transformation minimised likely variations in the speed of responding (Barnes-Holmes, et al., 2010). The transformation was conducted according to the following steps: 1. Only response latency data from test blocks was used. 2. Latencies >10,000 ms. were removed. 3. Data containing 10%+ of test trial-types with latencies <300 ms. were removed. 4. Twelve standard deviations for the four trial-types were calculated for each test block (four from Blocks 1 and 2, four from Blocks 3 and 4, and four from Blocks 5 and 6). 5. Twenty-four mean latencies were calculated for the four trial-types in each test block. 6. Difference scores for each trial-type were calculated for each pair of test blocks. This involved subtracting the mean latency of the consistent block from the mean latency of the corresponding inconsistent block. 7. Each difference score was divided by its corresponding standard deviation from Step 4. This generated 12 D_{IRAP} scores (one per trial-type per pair of test blocks). 8. Four overall trial-type D_{IRAP} scores were calculated. This involved averaging the three scores for each trial-type across the three pairs of test blocks. All data from any participant which fell

below 80% accuracy and above 2000ms latency on any of the six test blocks were omitted from the dataset ($N=24$). The final dataset comprised of 39 participants: 13 in each of the three IRAPs.

Between groups analyses. The mean D_{IRAP} scores for the three IRAPs are presented in Figure 2. On the Disorder-Negative trial-type, all groups showed stigmatising effects (i.e. faster responding on Disorder-Negative-True). However, on the Disorder-Positive trial-type, all groups showed non-stigmatising effects (i.e. faster responding on Disorder-Positive-True). On the Normality-Negative trial-type, the Anxiety IRAP showed pro-normality effects (i.e. faster responding on Normality-Negative-False), whereas the Mental Illness IRAP showed an anti-normality effect (i.e. faster responding on Normality-Negative-True). No effect was recorded on the Depression IRAP. On the Normality-Positive trial-type, all groups showed pro-normality effects (i.e. faster responding on Normality-Positive-True).

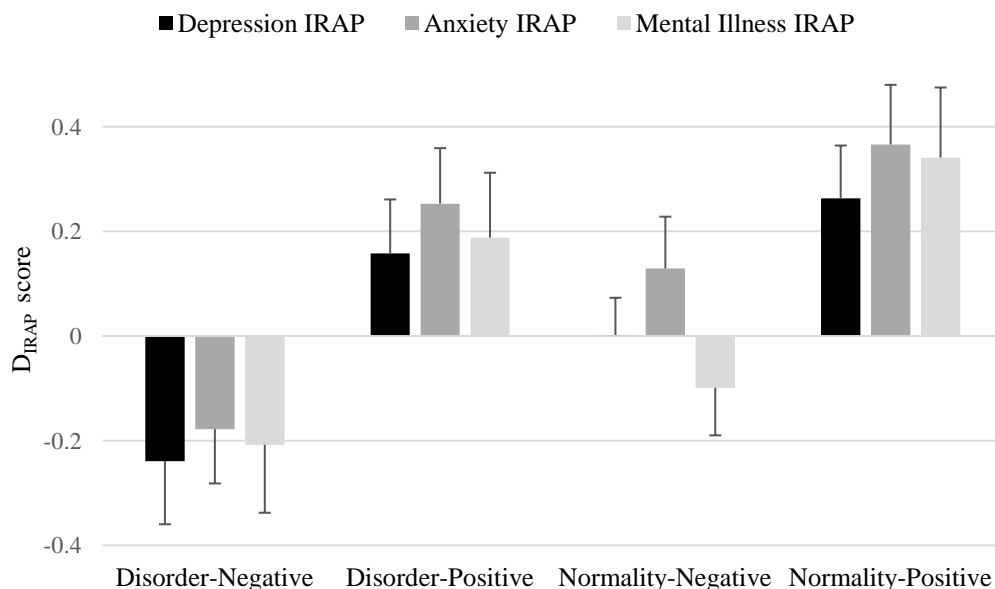


Figure 2. Mean D_{IRAP} scores for the three IRAPs in Experiment 1. **Positive D_{IRAP} scores indicate non-stigmatising/pro-normality effects and negative D_{IRAP} scores indicate stigmatising/anti-normality effects.**

In order to investigate the effects of IRAP on trial-type, a mixed between within ANOVA was conducted and a main effect for trial-type was recorded ($df=3$, $F=12.015$, $p<0.0001$, $\mu^2=1.0$). Post-hoc trial-type analyses, as 12 independent t-tests investigated potential differences across the IRAPs on each trial-type, but none were found (*all p's* >0.05). Subsequently, 12 one-sample t-tests investigated whether the D_{IRAP} trial-type effects differed significantly from zero. For each IRAP, the Normality-Positive trial-type was significant (Depression: $df=12$, $t=-2.602$, $p<0.05$; Anxiety: $df=12$, $t=-3.222$, $p<0.01$; Mental Illness: $df=12$, $t=-2.545$, $p<0.05$). And on the Anxiety IRAP, Disorder-Positive was also significant ($df=12$, $t=-2.386$, $p<0.05$).

Distress level analyses. Distress level analyses investigated the putative relationship between explicit level of distress and responding on the three IRAPs. For these analyses, the IRAP effects were divided along the clinical categories of the *overall DASS score* (i.e. Mild=0-42, Moderate=43-61, and Severe and Extreme=62+).

Depression IRAP. The mean D_{IRAP} scores for the three distress levels on the Depression IRAP are presented in Figure 3 (Mild, $N=9$, Moderate, $N=0$, Severe and Extreme, $N=4$). On Disorder-Negative, both groups showed stigmatising effects, the larger of which was in the mildly distressed group. On Disorder-Positive, both groups showed similar non-stigmatising effects. On Normality-Negative, the mildly distressed group showed anti-normality effects, whereas the severely distressed group showed pro-normality effects. On Normality-Positive, both groups showed pro-normality effects, the larger of which was in the severely distressed group.

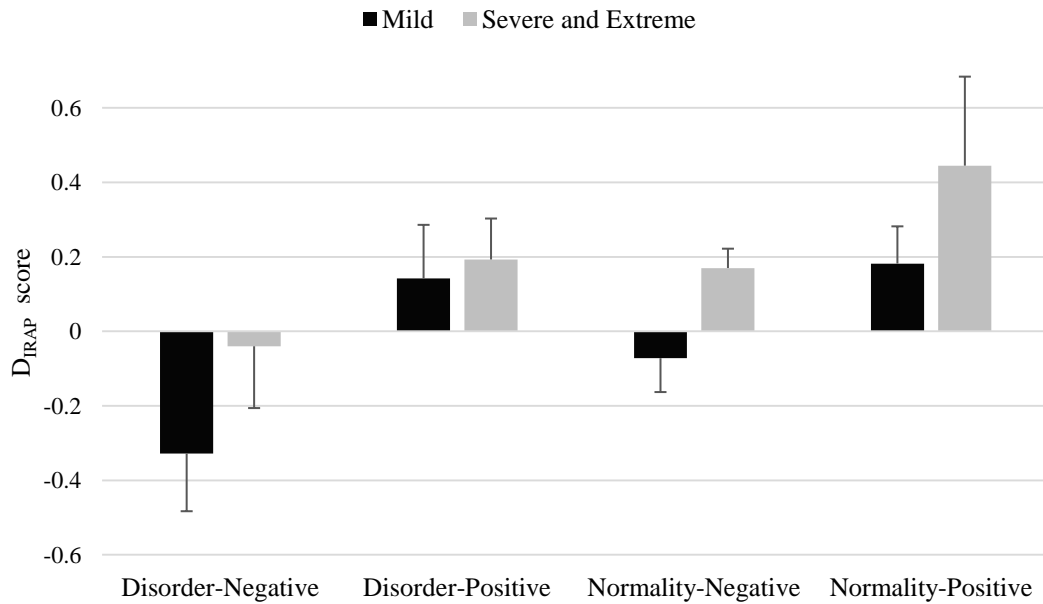


Figure 3. Mean D_{IRAP} scores for the two DASS distress groups on the Depression IRAP in Experiment 1. **Positive D_{IRAP} scores indicate non-stigmatising/pro-normality effects and negative D_{IRAP} scores indicate stigmatising/anti-normality effects.**

A mixed between within ANOVA investigated the effects of distress on IRAP trial-types, and found significant effects for trial type ($df=3$, $F=3.508$, $p<0.05$, $\mu^2=0.730$) and DASS group ($df=1$, $F=4.855$, $p<0.05$, $\mu^2=0.513$). Post-hoc trial-type analyses, in the form of four independent t-tests, were conducted, but no differences emerged (all $p's>0.05$). However, six one-sample t-tests showed that for the severely distressed group, Normality-Positive was significant ($df=3$, $t=-3.262$, $p<0.05$). All other effects were non-significant (all $p's >0.05$).

Anxiety IRAP. The mean D_{IRAP} scores for the three distress levels on the Anxiety IRAP are presented in Figure 4 (Mild, $N=5$, Moderate, $N=5$, Severe and Extreme, $N=3$). On Disorder-Negative, all groups showed stigmatising effects, the largest of which was in the severely distressed group. On Disorder-Positive, all groups showed non-stigmatising effects. On Normality-Negative, the mildly distressed group showed non-stigmatising effects. On Normality-Positive, the mildly distressed group showed pro-normality effects, whereas the severely distressed

groups showed anti-normality effects. No effect was recorded for the moderately distressed group on this trial-type. On Normality-Positive, all groups showed pro-normality effects, the largest of which was in the moderately distressed group.

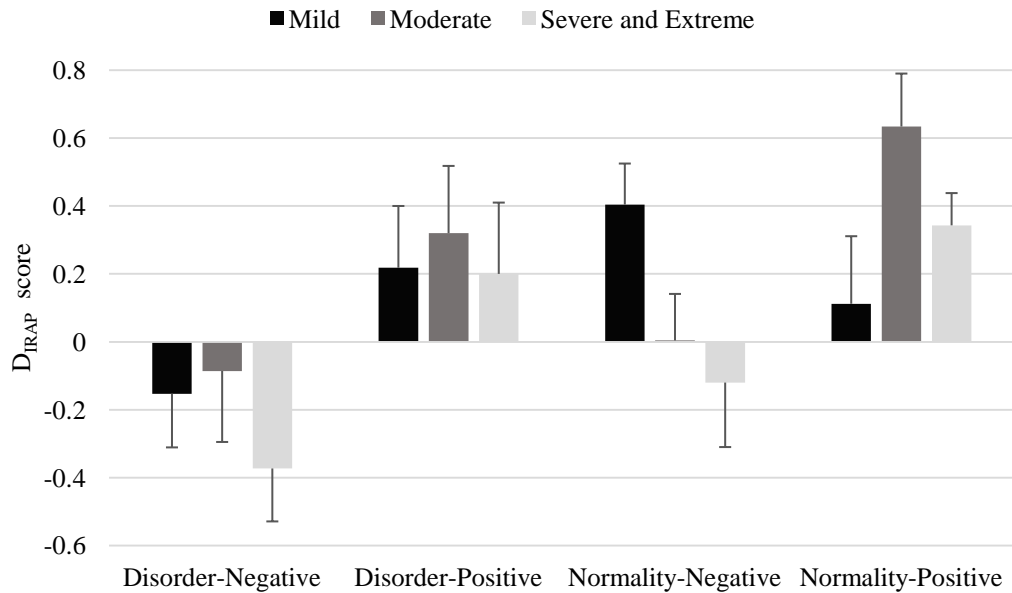


Figure 4. Mean D_{IRAP} scores for the three DASS distress groups on the Anxiety IRAP in Experiment 1. **Positive D_{IRAP} scores indicate non-stigmatising/pro-normality effects and negative D_{IRAP} scores indicate stigmatising/anti-normality effects.**

A mixed between within ANOVA investigated the effects of distress on IRAP trial-types and found a main effect for trial type ($df=3$, $F=4.573$, $p<0.01$, $\mu^2=0.848$). Post-hoc trial-type analyses (12 independent t-tests) found differences on Normality-Bad between the mildly and severely distressed groups ($df=6$, $t=-2.462$, $p<0.05$), while the difference between the mildly and moderately distressed groups approached significance ($df=8$, $t=-2.185$, $p=0.06$). All other differences were not significant (all $p's>0.05$). However, six one-sample t-tests showed that for the mildly distressed group, the D_{IRAP} effect was significant for Normality-Negative ($df=4$, $t=-3.342$, $p<0.05$). For the moderately distressed group, the D_{IRAP} effect was significant

for Normality-Positive ($df=4$, $t=-4.075$, $p<0.05$). All other effects were non-significant (all p 's >0.05).

Mental illness IRAP. The mean D_{IRAP} scores for the three distress levels on the Mental Illness IRAP are presented in Figure 5 (Mild, $N=8$, Moderate, $N=2$, Severe and Extreme, $N=3$). On Disorder-Negative, all groups showed stigmatising effects, the largest of which was in the mildly distressed group. However, Figure 5 shows considerable variance on this trial-type in the moderately distressed group (as depicted by the standard error bar). On Disorder-Positive, the mildly and severely distressed groups showed non-stigmatising effects, whereas the moderately distressed group showed marginal stigmatising effects. On Normality-Negative, all groups showed anti-normality effects, the largest of which was in the moderately distressed group. On Normality-Positive, all groups showed pro-normality effects.

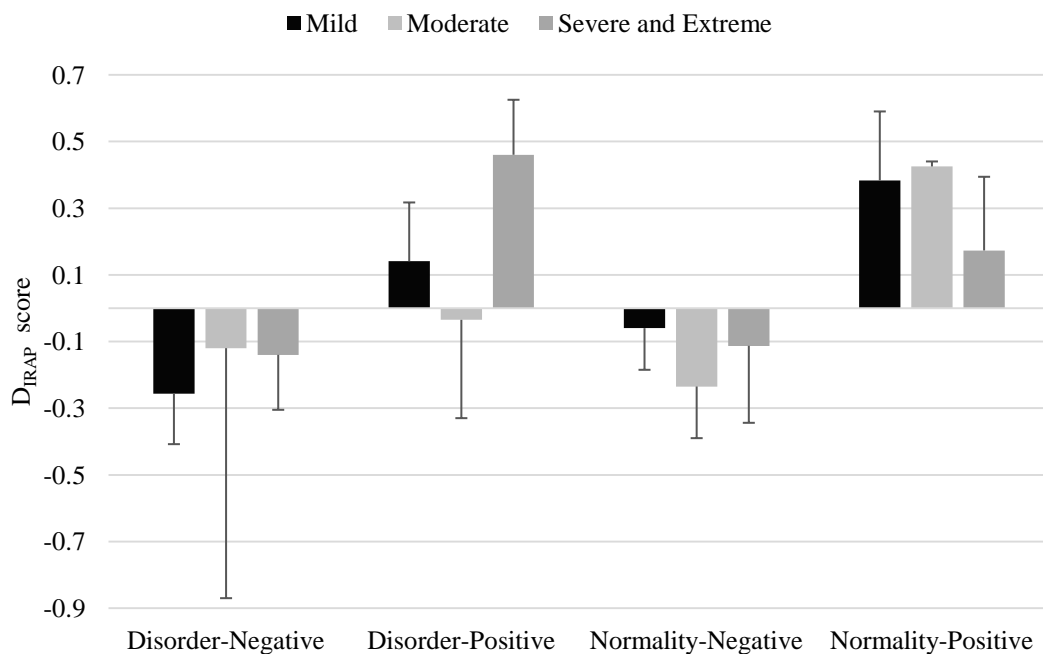


Figure 5. Mean D_{IRAP} scores for the three DASS distress groups on the Mental Illness IRAP in Experiment 1. **Positive D_{IRAP} scores indicate non-stigmatising/pro-normality effects and negative D_{IRAP} scores indicate stigmatising/anti-normality effects.**

A mixed between within ANOVA investigated the effects of distress on IRAP trial-types but none was found (all p 's > 0.05). Exploratory trial-type analyses (12 independent t-tests) found no differences among the distress groups (all p 's > 0.05). However, six one-sample t-tests showed that for the moderately distressed group, the D_{IRAP} effect was significant for Normality-Positive ($df=1$, $t=-28.333$, $p<0.05$). However, this effect is based on a very small N and therefore should be treated with caution. All other effects were non-significant (all p 's > 0.05).

Correlations

A correlation matrix was conducted between the IRAP trial-types and the explicit measures. For the Depression IRAP, Disorder-Negative showed significant negative correlations with DASS Anxiety ($df=13$, $r=-0.568$, $p<0.05$). That is, as anxiety decreases, stigmatising effects increase. For Disorder-Positive, significant negative correlations were found with CAMI Authoritarianism ($df=13$, $r=0.688$, $p<0.01$) and Social Restrictiveness ($df=13$, $r=-0.563$, $p<0.05$). That is, as stigma decreases, non-stigmatising IRAP effects increase. Furthermore, significant negative correlations were also found with CAMI Benevolence ($df=13$, $r=0.770$, $p<0.01$) and CAMI Mental Health Ideology ($df=13$, $r=0.667$, $p<0.05$). That is, as acceptance of mental health issues increase, so too do non-stigmatising effects. For Normality-Positive, a significant negative correlation was found with DASS depression ($df=13$, $r=0.537$, $p<0.05$). That is, as depression and overall distress increase, pro-normality effects also increase.

For the Anxiety IRAP, Normality-Negative showed significant positive correlations with Total DASS ($df=13$, $r=0.718$, $p<0.01$) and Anxiety ($df=13$,

$r=0.667$, $p<0.01$). That is, as anxiety and overall distress increase, pro-normality effects also increase.

Bonferroni adjustment. Due to multiple correlational analyses conducted, a Bonferroni adjustment of the alpha level was applied ($p<0.0005$). However, using this p-value, the above correlations do not reach significance (all $p's>0.0005$). Indeed, while this adjustment is deemed necessary for statistical conservatism and to avoid the occurrence of a Type 1 error, for the purposes of the current thesis, the correlations which were significant at the 0.05 level will also be discussed.

Summary and Discussion

Experiment 1 assessed evaluations of (and stigmatisation to) psychological suffering, by juxtaposing suffering with normality using positive versus negative comparisons. Each of three IRAPs targeted a specific type of suffering, namely depression (using the Depression IRAP), anxiety (using the Anxiety IRAP) and the concept of mental illness generally (using the Mental Illness IRAP). Overall, all three IRAPs produced similar patterns of responding, which, surprisingly, indicated that all three forms of suffering were implicitly evaluated as **both** positive (e.g. anxious people are-relaxed-true) *and* negative (e.g. depressed people are-lazy-true). Similarly, all three IRAPs also showed significant pro-normality effects (e.g. normal people are-happy-true), although an anti-normality effect (e.g. normal people are-sad-true) was found on the Mental Illness IRAP.

In spite of this consistency of effects across the three IRAPs, the distress level analyses showed more divergent outcomes. On the Depression IRAP, mildly distressed participants evaluated depression most negatively, while the severely distressed evaluated normality most positively, and both of these effects were

significant. On the Anxiety IRAP, the severely distressed evaluated both anxiety and normality most negatively, whereas the mildly and moderately significantly evaluated normality most positively. On the Mental Illness IRAP, there was general negativity toward mental illness, but only the mildly and severely distressed were also positive in this regard. Taken together, these effects suggest that there is a relationship between psychological suffering and evaluations of suffering and normality, but these effects are clearly subtle and complex.

The correlations revealed that some IRAP trial-types predicted some aspects of explicit. For example, on the Depression IRAP, positivity toward depression correlated with low anxiety and low stigma, while negativity correlated with stigma. Positivity toward normality correlated with high distress. For the Anxiety IRAP, pro-normality effects correlated with overall distress. Overall, these correlations suggest that psychological distress impacts upon implicit evaluations to suffering and normality, and thus, may influence stigmatisation.

It was somewhat interesting to find that the participants from Experiment 1 did not implicitly evaluate depression, anxiety and mental illness as negatively as one would expect, and thus were not particularly stigmatising toward these concepts. This finding appears to contrast the findings of Lincoln et al. (2008) where 'schizophrenia' was stigmatised the most. However, the current research did use the more generic concept of 'mental illness', which may indeed be less stigmatising than the concept of 'schizophrenia'. However, it was not surprising of course, to find that this sample had strong implicit positivity toward normality. However, the analyses also showed that there is likely to be a relationship between one's own level of distress and the nature and extent to which one evaluates psychological suffering and normality, which supports research conducted by Rüsçh et al. (2010), where self-

stigma predicted a lower quality of life. These broad but nuanced findings provided an important baseline for the subsequent studies in terms of allowing us to benchmark IRAP outcomes for very specific psychological phenomena against outcomes on broad concepts such as depression, anxiety and mental illness. In short, given that we now have some indication of how participants evaluate normality and psychological suffering broadly, we are now in a better position to determine how they will evaluate more specific concepts of psychological suffering, relatively speaking. Although, it is important to recognise that there were very few significant IRAP effects in Experiment 1, the various analyses and findings suggested that the procedure would provide us with a useful methodology for further investigations in the target phenomena. These more specific investigations began in Experiment 2.

Limitations

There are a number of limitations of Experiment 1 which should be reflected on to guide future research. 1. The use of inferential statistics in low N analyses (i.e. the distress level analyses) does not allow researchers to observe the differences or effects which may exist. Future research should try to include larger N 's in the distress level analyses in order to better examine this nuanced relationship between one's own level of distress and implicit attitudes. 2. The stimuli used in the three IRAPs may not fully capture the negative connotations of stigma, and may simply reflect the patterns of traits often observed in each form of suffering. Future studies could include stimuli which target stigmatising *behaviours* toward sufferers (e.g. I would stay away from these people, they are different to me etc.). 3. The concept of normality was selected in this study as a contrast category for suffering, however, in reality 'normality' (i.e. normal verbal behaviour) also encompasses aspects of

suffering. Future studies could investigate the use of alternative contrast categories which adequately encapsulate the absence of specific forms of suffering or diagnoses. 4. The current sample comprised of a high proportion of psychology undergraduates exposed to some level of psychological training, which may account for some of the positivity, therefore, it would be interesting to replicate this study in a sample with no training in psychology.

Experiment 2

Assessing Implicit Locus of Control in Psychological Suffering

The literature suggests that the concept of locus of control may play a role in stigmatising attitudes toward psychological suffering (Rotter, 1966). One of the few IAT studies on stigma toward psychological suffering reported by Teachman and colleagues (2006) investigated stigma against suffering versus physical illness. Experiment 2 followed this line of inquiry to investigate the locus of control in psychological suffering versus illness. That is, we developed two IRAPs, namely an Illness IRAP and a Weakness IRAP to investigate internal versus external locus of control. If it was the case that participants perceived physical illness as a source of external locus of control, but perceived weak-mindedness as an internal locus of control, then we would hypothesise more positive evaluations to illness and more negative evaluations to weakness.

Method

Setting

All aspects of the setting in Experiment 2 were identical to Experiment 1.

Participants

The current study involved undergraduate students recruited from MU. Sixty-four individuals participated, 25 were female and 39 were male, with an age range of 18 to 23 years and a mean of age of 20.5 years and a standard deviation of 2.44 years.

Materials

Explicit measures. Three broad categories of explicit measures were administered. The first series of measures assessed stigma and locus of control (the CAMI, the SAB and the Illness Questionnaire). The second measure more broadly assessed general psychological well-being (the DASS). See Experiment 1 for outlines of the CAMI, SAB and DASS.

The Illness Questionnaire (IQ). The IQ was constructed for the purpose of this study. This 6-item explicit measure determines whether participants believe that various mental health conditions (i.e. depression, anxiety, phobia, worry, obsession and panic) are illness-based (e.g. “do you think depression is an illness?”). Participants responded to various questions by circling “Yes” or “No”. Scores ranged from 0-6. A score of 6 indicated participants believed that the various mental health conditions were illness-based and a score of 0 indicated that participants did not believe the various mental health conditions were illness-based (See Appendix 5). However, the internal consistency for the measure is deemed statistically questionable ($\alpha=0.603$).

The IRAP. The IRAP was administered on a standard desktop computer. This was used to present the instructions and stimuli, and to record responses. The current study involved two IRAPs which assessed positive or negative evaluations of psychological suffering as illness and weak-mindedness (referred to as the Illness IRAP and the Weakness IRAP). The IRAP stimuli were adapted from the three IRAPs in Experiment 1 and aim to represent some of the broad features of psychological suffering juxtaposed against psychological health. The sample stimuli ILL PEOPLE ARE and WEAK-MINDED PEOPLE ARE contrast the locus as

external (as a physical illness) or internal (as a psychological weakness) using the same target stimuli in each IRAP to allow a clear distinction between the loci only.

The IRAPs contrasted illness and weak-mindedness with health, using the labels ILL PEOPLE ARE and HEALTHY PEOPLE ARE. Each trial-type presented one of these two types of category labels. These were accompanied by one of six positive (e.g. HAPPY) or six negative target stimuli (e.g. DEPRESSED). The same target stimuli were used in both the Illness and the Weakness IRAPs. The screen also presented two response options, TRUE and FALSE. Before each block of trials, the screen presented one of two rules for responding (e.g. PLEASE ANSWER AS IF ILL PEOPLE ARE NEGATIVE AND HEALTHY PEOPLE ARE POSITIVE or PLEASE ANSWER AS IF HEALTHY PEOPLE ARE NEGATIVE AND DEPRESSED PEOPLE ARE POSITIVE). A full list of label stimuli, target stimuli and response options for the IRAPs is provided in Table 3.

Table 3

Stimuli and Response Options of the two IRAPs employed in Experiment 2

Illness IRAP		Weakness IRAP	
Ill People are	Healthy People are	Weak-minded People are	Healthy People are
Phobic	Rational	Phobic	Rational
Depressed	Happy	Depressed	Happy
Anxious	Relaxed	Anxious	Relaxed
Worried	Mellow	Worried	Mellow
Panicky	Calm	Panicky	Calm
Obsessive	Easy Going	Obsessive	Easy Going
True	False	True	False

Procedure

The current study comprised of two IRAPs: The Illness IRAP and The Weakness IRAP. For all participants, there were two stages, one involving the implicit measure (an IRAP) and the second involving the explicit measures². It is important to emphasise, therefore, that each participant only completed *one IRAP*, the Illness IRAP *or* the Weakness IRAP, hence, approximately half of the total group of participants completed each IRAP. Thus, the design of the study was between groups repeated measures. Participants were quasi-randomly assigned to either IRAP and all participants completed the IRAP prior to the explicit measures.

The verbal and automated instructions provided to participants for completing the two IRAPs were identical to Experiment 1. For illustrative purposes, see Figure 6 for a schematic representation of the screen presentation of the IRAPs.

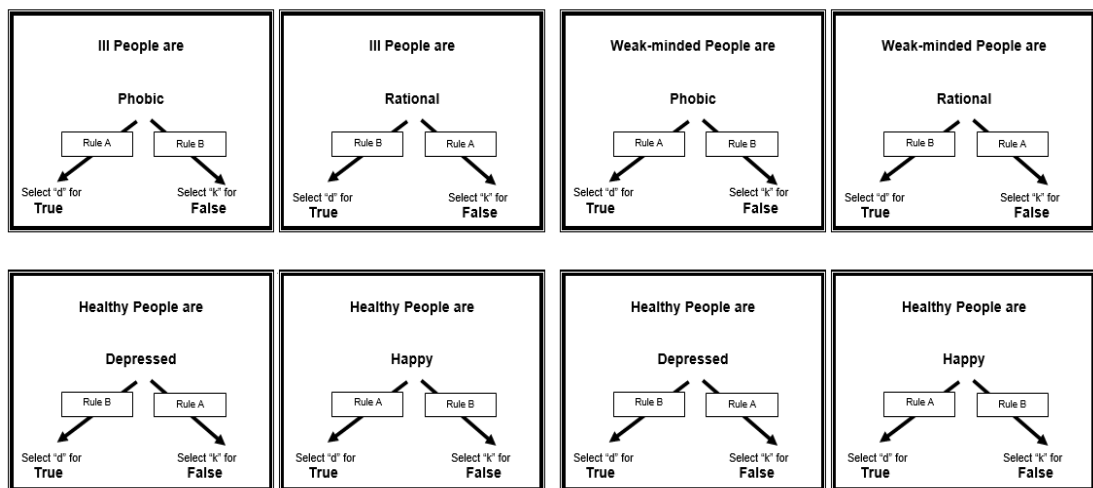


Figure 6. Schematic representation of the four trial-types presented in the Illness IRAP (left) and the Weakness IRAP (right) in Experiment 2. The arrows and text boxes did not appear on the participant’s screen, they indicate the correct responses for Rule A and Rule B blocks of trials. **The labels used for the four trial-types in**

² Notably, previous IRAP unpublished IRAP research has demonstrated that the order of presentation of the self-report measure does not influence IRAP results. Thus, to reduce the impact of fatigue following the completion of a number of explicit measures on IRAP performance, all participants completed the IRAP first.

each IRAP are as follows: Disorder-Negative (top-left), Disorder-Positive (top-right), Healthy-Bad (bottom-left) and Healthy-Good (bottom-right).

Once participants finished the IRAP, they completed the three explicit measures in a pre-determined sequence (CAMI, SAB, IQ and DASS).

Ethical Considerations

All aspects of potential ethical considerations in Experiment 2 were identical to Experiment 1.

Results

Analytic Strategy

To assess the potential effects of IRAP type on trial-type, a mixed between within ANOVA was conducted. Exploratory distress level analyses investigated potential influence of distress on stigmatisation on each IRAP. Hence, three mixed between within ANOVAs were conducted. Correlational analyses were then carried out to investigate the potential predictive validity of the IRAP in this context.

Before conducting the IRAP analyses, summaries of the explicit data are provided in the section below.

Explicit Measures Data

The means and standard deviations were calculated from each participant's responses on each of the explicit measures and data are summarised in Table 4.

Table 4

Explicit Data Summary for Experiment 2

Scales	Illness IRAP	Weakness IRAP	T-Test
	Mean (SD)	Mean (SD)	T
DASS	38.70 (27.38)	38.20 (26.00)	.059
Depression	11.00 (8.70)	9.20 (9.99)	.608
Anxiety	6.60 (6.90)	9.00 (7.66)	-1.041
Stress	20.80 (15.61)	20.00 (14.04)	.170
IQ	2.35 (1.42)	2.90 (1.83)	-1.060
CAMI	-41.95 (1.42)	-42.15 (17.37)	.041
Authoritarian	19.10 (13.17)	19.95 (4.30)	-.733
Benevolence	41.60 (4.29)	42.15 (5.62)	-.348
Social Restrictiveness	18.70 (4.03)	18.70 (6.80)	0
Mental health ideology	38.15 (5.74)	38.95 (4.70)	-.482
SAB Total	19.40 (5.46)	20.00 (5.92)	-.305

*Note. Maximum scores are: DASS total = 126; DASS subscales = 42; CAMI = 10-50; SAB = 56. Significance indicated by *($p < 0.05$).

For the DASS, means revealed that participants overall had low rates of depression, anxiety and stress. For the CAMI, means revealed that participants overall had low stigmatisation and high acceptance of mental illness. And for the SAB, means revealed low stigmatising attitudes toward mental illness. Participants' outcomes on the IQ demonstrated a weak-moderate belief across both IRAPs that psychological suffering was illness-based. Notably, no differences existed between the explicit data across the two IRAPs using independent t-tests (all p 's > 0.05).

IRAP Data

Scoring of the IRAP was always conducted using the standardised approach for transforming latency data into D_{IRAP} scores (see Nicholson & Barnes-Holmes, 2012a). All data from any participant which fell below 80% accuracy and above 2000ms latency on any of the six test blocks were omitted from the dataset ($N=24$). The final dataset comprised of 40 participants: 20 participants in each IRAP.

Between groups analyses. The mean D_{IRAP} scores for the two IRAPs are presented in Figure 7. On the Disorder-Negative trial-type, both groups showed stigmatising effects, the larger of which was found in the Illness IRAP. However, on the Disorder-Positive trial-type, participants showed marginal non-stigmatising effects. On the Healthy-Bad trial-type, both groups showed pro-healthy effects, the larger of which was found in the Weakness IRAP. Also, on the Healthy-Good trial-type, both groups showed comparable pro-normality effects.

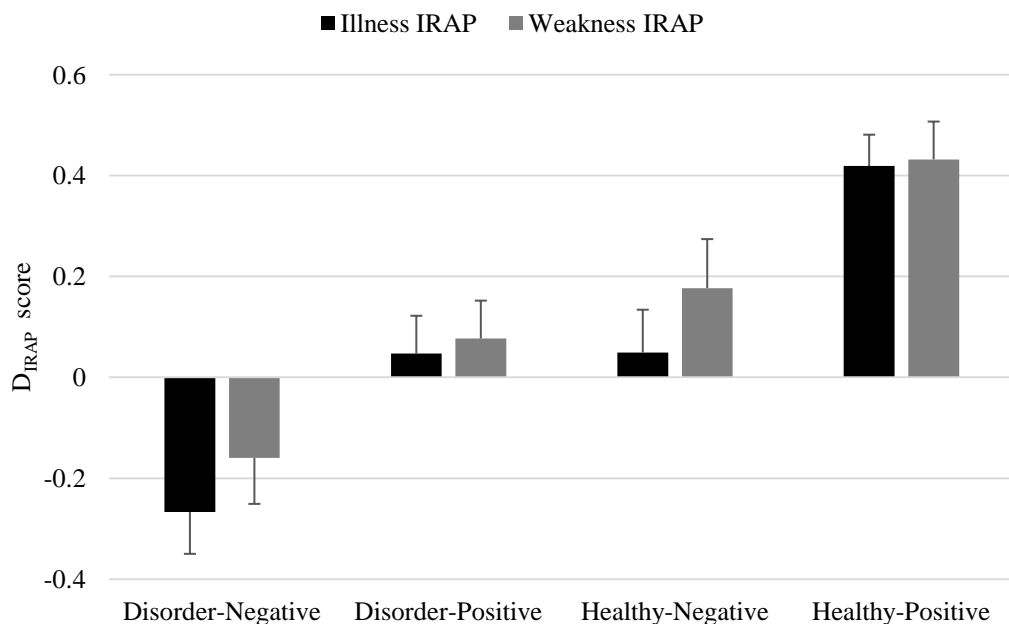


Figure 7. Mean D_{IRAP} scores for the two IRAPs in Experiment 2. **Positive D_{IRAP} scores indicate non-stigmatising/pro-healthy effects and negative D_{IRAP} scores indicate stigmatising/anti-healthy effects.**

In order to investigate the effects of group on trial-type, a mixed between within ANOVA was conducted and a main effect for trial-type was found ($df=3$, $F=23.257$, $p<0.0001$, $\mu^2=1.0$). Post-hoc trial-type analyses, as four independent t-tests, investigated potential differences between IRAPs on each trial-type, but no differences were found (all $p's>0.05$). Eight one-sample t-tests investigated whether

the D_{IRAP} trial-type effects differed significantly from zero. Healthy-Good was significant for both IRAPs (Illness: $df=19$, $t=-6.720$, $p<0.0001$; Weakness: $df=19$, $t=-5.752$, $p<0.0001$). Disorder-Negative was also significant for the Illness IRAP ($df=19$, $t=3.204$, $p<0.01$).

Distress level analyses. Distress level analyses investigated the putative relationship between explicit level of distress and responding on the two IRAPs. For these analyses, the groups were again divided along the clinical categories of the overall DASS score (i.e. Mild, Moderate, and Severe and Extreme).

Illness IRAP. The mean D_{IRAP} scores for the three distress levels on the Illness IRAP are presented in Figure 8 (Mild, $N=14$, Moderate, $N=3$, Severe and Extreme, $N=3$). On Disorder-Negative, the mild and extreme groups showed stigmatising effects, but the moderately distressed group showed non-stigmatising effects. Similarly, on Disorder-Positive, the mild and extreme groups showed non-stigmatising effects, but the moderately distressed group showed stigmatising effects. On Healthy-Bad, the moderate and severe groups showed anti-healthy effects, whereas the mild group showed pro-healthy effects. On Healthy-Good, all groups showed pro-normality effects, the largest of which was in the mild group.

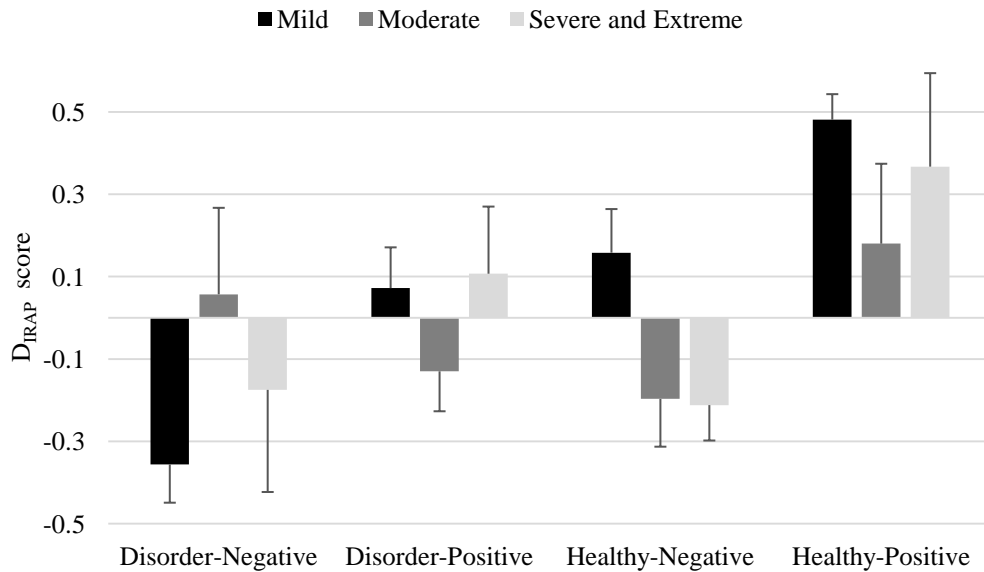


Figure 8. Mean D_{IRAP} scores for the three DASS distress groups on the Illness IRAP in Experiment 2. **Positive D_{IRAP} scores indicate non-stigmatising/pro-healthy effects and negative D_{IRAP} scores indicate stigmatising/anti-healthy effects.**

A mixed between within ANOVA investigated the effects of distress on IRAP trial-types, and found a significant effect for trial-type ($df=3$, $F=7.083$, $p<0.001$, $\mu^2=0.980$) and an interaction effect ($df=6$, $F=2.396$, $p<0.05$, $\mu^2=0.769$). However, post-hoc trial-type analyses (12 independent t-tests) found no significant differences (all $p's>0.05$). Nonetheless, 12 one-sample t-tests showed that for the *mildly* distressed group, the D_{IRAP} effect was significant for Disorder-Negative ($df=13$, $t=3.852$, $p<0.01$) and Healthy-Good ($df=13$, $t=-7.716$, $p<0.0001$).

Weakness IRAP. The mean D_{IRAP} scores for the three distress levels on the Weakness IRAP are presented in Figure 9 (Mild, $N=13$, Moderate, $N=2$, Severe and Extreme, $N=5$). On Disorder-Negative, the mild group showed stigmatising effects, whereas the moderate and severe groups (marginally for the latter) showed non-stigmatising effects. On Disorder-Positive, all groups showed non-stigmatising effects. On Healthy-Bad and Healthy-Good, all groups showed pro-normality effects.

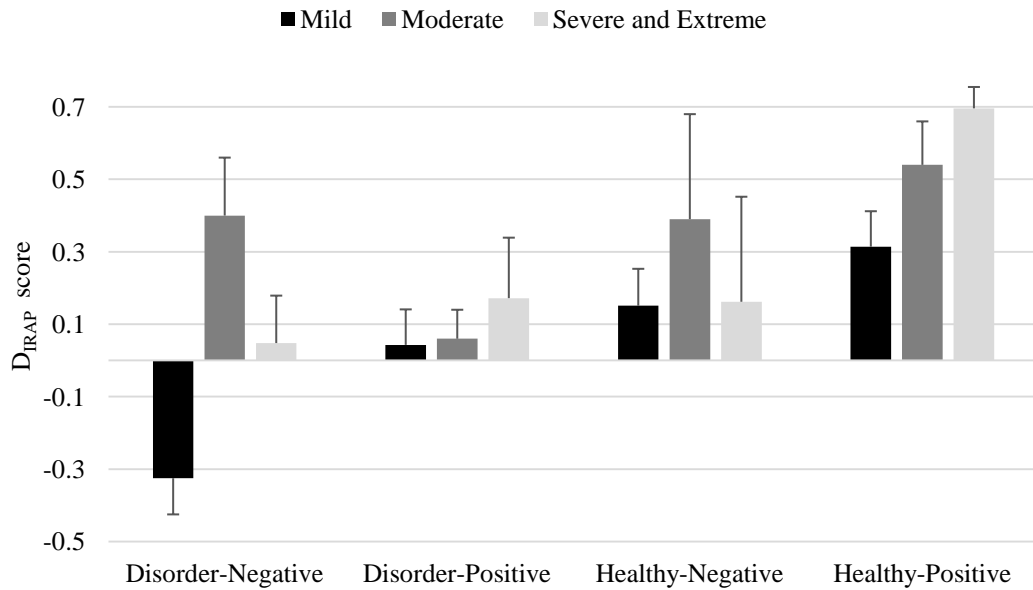


Figure 9. Mean D_{IRAP} scores for the three DASS distress groups on the Weakness IRAP in Experiment 2. **Positive D_{IRAP} scores indicate non-stigmatising/pro-healthy effects and negative D_{IRAP} scores indicate stigmatising/anti-healthy effects.**

A mixed between within ANOVA investigated the effects of distress on IRAP trial-types, and found significant effects for trial-type ($df=3$, $F=3.749$, $p<0.05$, $\mu^2=0.783$) and distress group ($df=2$, $F=5.125$, $p<0.05$, $\mu^2=0.751$). Post-hoc trial-type analyses (12 independent t-tests) found that on Disorder-Negative, the mild and moderate groups differed significantly ($df=13$, $t=2.707$, $p<0.05$). On Healthy-Good, the mild and severe groups differed significantly ($df=16$, $t=2.325$, $p<0.05$). Furthermore, 12 one-sample t-tests showed that for the mild group, the D_{IRAP} effect was significant for Disorder-Negative ($df=12$, $t=3.247$, $p<0.01$) and Healthy-Good ($df=12$, $t=-3.209$, $p<0.01$). For the severe group, the Healthy-Good trial-type was also significant ($df=4$, $t=-11.822$, $p<0.001$).

Correlations

A correlation matrix was conducted between the IRAP trial-types and the explicit measures. For the Illness IRAP, a significant negative correlation was found between Healthy-Bad and CAMI Authoritarianism ($df=20$, $r=-0.439$, $p<0.05$). That is, the greater the stigma, the greater the pro-healthy effects on the IRAP.

For the Weakness IRAP, significant negative correlations were found for Disorder-Negative with Total DASS ($df=20$, $r=-0.574$, $p<0.01$) and stress ($df=20$, $r=-0.623$, $p<0.01$). That is, the greater the stigmatising effect, the less overall suffering on the DASS, and the greater the pro-healthy effects, the greater the overall suffering on the DASS. Significant positive correlations were also found for Healthy-Good and SAB ($df=20$, $r=0.522$, $p<0.05$), suggesting that the greater the pro-healthy bias, the greater the stigma.

Bonferroni adjustment. Due to multiple correlational analyses conducted, a Bonferroni adjustment of the alpha level was applied ($p<0.0005$). However, using this p-value, the above correlations do not reach significance (all $p's>0.0005$). Indeed, while this adjustment is deemed necessary for statistical conservatism and to avoid the occurrence of a Type 1 error, for the purposes of the current thesis, the correlations which were significant at the 0.05 level will also be discussed.

Summary and Discussion

The two IRAPs employed in Experiment 2 assessed perceptions of the locus of control for, and evaluations of, psychological suffering by juxtaposing suffering with health using positive versus negative comparisons. Each IRAP targeted suffering as either an illness (using the Illness IRAP) or as weak-mindedness (using the Weakness IRAP). Both IRAPs produced similar patterns of responding, in which

suffering (as illness or weakness) was evaluated as both positive (e.g. ill people are-happy-true) *and* negative (e.g. ill people are-depressed-true). Significantly strong positivity toward health was also recorded.

Again, the distress level analyses generated divergent effects. Mildly distressed participants evaluated illness and weakness most negatively on both IRAPs, and both of these effects were significant. This group also evaluated health most positively on the Illness IRAP, but it was the severely distressed who evaluated health most positively on the Weakness IRAP, and these effects were also significant. The correlations revealed that positivity toward health correlated with high stigma on the Illness IRAP and high distress and stigma on the Weakness IRAP, on which negativity toward weakness also correlated with low distress.

These findings bore considerable overlap with those from Experiment 1 in terms of showing less implicit negativity toward weak-mindedness than one might traditionally expect, although again there was considerable positivity toward health. It was interesting that negativity towards illness was only marginally greater than toward weakness, and indeed the lack of difference across the two IRAPs suggests that participants overall do not attribute the locus *differentially* in physical illness or weak-mindedness, which contrasts with effects reported by Teachman et al. (2006), who found no stigmatising effects for physical illness. Once again, these evaluations appear to be mediated by a participant's own level of distress, and in short, participants who were mildly distressed were the most negative towards illness and weak-mindedness. These analyses begin to suggest the robustness of the patterns, albeit non-significant for the most part, observed in the IRAPs in Experiment 1. Furthermore, we have now an even broader baseline understanding of participants' implicit evaluations and stigma regarding psychological suffering. This now

provided a stronger springboard for the more specific analyses of implicit attitudes to voice hearing that were the focus of all subsequent experiments in the current thesis. This analysis commenced in Experiment 3.

Limitations

There are a number of limitations of Experiment 2 which should be reflected on to guide future research. 1. The use of inferential statistics in low N analyses (i.e. the distress level analyses) does not allow researchers to observe the differences or effects which may exist. Future research should try to include larger N 's in the distress level analyses in order to better examine this nuanced relationship between one's own level of distress and implicit attitudes. 2. The use of binary response options in the IQ may reduce the sensitivity of the measure. Future studies should use a Likert scale of measurement. 3. The concept of health was selected in this study as a contrast category for illness or weak-mindedness, however, future studies could investigate the use of alternative contrast categories which may avoid the use of the health as a contrast category (i.e. physical illness versus psychological suffering). Target stimuli selection could also be aided through the use of illness explanatory models (Kleinman, Eisenberg, & Good, 1978). 4. The current sample comprised of a high proportion of psychology undergraduates exposed to some level of psychological training, therefore, it would be interesting to replicate this study in a sample with no training in psychology.

Chapter 3
Experiment 3

Assessing evaluations of voice
hearing in a non-voice hearing
sample

Experiment 3

Assessing Implicit Evaluations of Voice Hearing

All published studies of voice hearing have relied solely on explicit measures (Kim et al., 2010). Hence, the literature contains no single study using implicit measures to investigate this phenomenon. Furthermore, there are very few studies that have assessed evaluations of voice hearing in participants who do not report hearing voices. Without this type of data, it is difficult to determine whether evaluations shown by voice hearers are not also shared by participants and thus are not specific to individuals who have direct experiences in this regard. Toward this aim, Experiment 3 sought to investigate stigmatisation of hearing voices by a student population. The study employed a simple Evaluation IRAP which juxtaposed hearing voices with seeing things in the context of positive and negative evaluations. It must be emphasised that the contrast category of seeing things was selected for purely experimental reasons because it is very difficult to generate relevant categories about hearing voices for individuals who have never had this experience. Nonetheless, a focus on seeing things offered an interesting relative comparison for hearing voices. In such an exploratory study, we hypothesised that hearing voices would be evaluated negatively, however, it was difficult to make predictions regarding IRAP effects, especially when participants were being presented with concepts with which they have no direct experience. Hence, our simple aim was to determine whether IRAP effects would be observed in this context and to what extent these might correlate with explicit measures, as they had done in the previous two studies.

Method

Setting

All aspects of the setting in Experiment 3 were identical to Experiment 1.

Participants

The current study involved a group of non-voice hearing participants who were identified as such using current screening methods from a general sample of undergraduate students recruited from MU. Thirty-five non-voice hearing individuals were identified, 12 of these were male and 23 were female, with an age range of 18 to 26 years and a mean age of 21.72 years and a standard deviation of 1.55 years.

Materials

Explicit measures. Three broad categories of explicit measures were administered. The first series of measures assessed voice hearing (the AHRS) and psychosis (CAPE). The second set of measures more broadly assessed general psychological well-being (the AAQ-II and the ATQ). The third measured stigma toward mental health difficulties (SAB, see Chapter 2).

Auditory Hallucinations Ratings Scale (AHRS; Haddock et al., 1999). The AHRS is a subscale of the Psychotic Symptom Rating Scales (PSYRATS) and excludes an additional subscale that measures delusions. The AHRS is an 11-item scale that assesses the severity of 11 target dimensions of voice hearing (e.g. degree of negative content - minority of voice content is unpleasant or negative). All items are rated on a 5-point scale from 0 (e.g. voices not present) to 4 (e.g. voices present most of the time). The AHRS yields an overall score with a maximum of 44

indicating high degrees of voice hearing and a minimum of 0 indicating low degrees of voice hearing. This scale has demonstrated excellent inter-rater reliability (ICC = 0.99 - 1.00) and test-retest reliability (ICC = 0.70; Drake, Haddock, Tarrier, Bentall, & Lewis, 2007; Haddock et al., 1999). See Appendix 6.

Community Assessment of Psychic Experience (CAPE; Stefanis et al., 2002). The CAPE is a 42-item measure of general delusional ideation (derived from the Peters Delusions Inventory, PDI, Peters, Joseph, & Garety, 1999). The measure assesses three dimensions of psychotic-like experiences: positive, negative and depressive (e.g. “do you ever feel as if a double has taken the place of a family member, friend or acquaintance?”). All items are rated in terms of frequency on a 4-point scale from 0 (never) to 1 (nearly always) and similarly rated in terms of level of distress from 0 (not distressed) to 3 (very distressed). These two sets of ratings on each dimension yield six independent scores (i.e. two scores for each dimension) that indicate high or low frequency or distress on each dimension. This scale has demonstrated adequate reliability with an alpha coefficient of 0.63 for the positive dimension, 0.64 for the negative dimension and 0.62 for the depressive dimension (Konings, Bak, Hanssen, Van Os, & Krabbendam, 2006). See Appendix 7.

Acceptance and Action Questionnaire II (AAQ-II; Bond et al., 2011). The AAQ-II is a 10-item measure of psychological inflexibility or avoidance (e.g. “my painful memories prevent me from having a fulfilling life”). All items are rated on a 7-point scale from 1 (never true) to 7 (always true). The AAQ-II yields an overall score with a maximum of 70 indicating *high* inflexibility and a minimum of 10 indicating *low* inflexibility. This scale has demonstrated adequate internal consistency with an alpha coefficient of 0.78 to 0.88 across several samples (Bond et al.). See Appendix 8.

Automatic Thoughts Questionnaire (ATQ; Hollon & Kendall, 1980). The ATQ is a 30-item measure of the frequency and believability of negative thoughts (e.g. “I feel like I’m up against the world”). The measure assesses four thought dimensions: personal maladjustment and desire for change (PMDC); negative self-concept and negative expectations (NSNE); low self-esteem (LSE); and helplessness (H). All items are rated in terms of frequency on a 5-point scale from 1 (never) to 5 (all the time) and similarly rated in terms of level of believability from 1 (not at all believable) to 5 (entirely). These two sets of ratings on each dimension yield eight independent scores (i.e. two scores for each dimension) that indicate high or low frequency or believability on each dimension. This scale has demonstrated excellent internal consistency with an alpha coefficient of 0.96 (Hollon & Kendall). See Appendix 9.

The IRAP. The IRAP was administered on a standard desktop computer. This was used to present the instructions and stimuli, and to record responses. The current study involved one IRAP which assessed positive or negative evaluations of voices (referred to as the Evaluation IRAP). The stimuli of the IRAP were selected to represent some of the broad evaluations of hallucinations. Notably, the positive stimuli were intentionally selected to be more ambivalent than highly positive, as we hypothesised that this would be unlikely in non-voice hearing participants who have never had these experiences.

The IRAP contrasted hearing voices with seeing things, using the labels HEARING VOICES IS and SEEING THINGS IS. Each trial-type presented one of these two types of category labels. These were accompanied by one of three positive (e.g. SAFE) or three negative target stimuli (e.g. BAD). The screen also presented two response options, TRUE and FALSE. Before each block of trials, the screen

presented one of two rules for responding (i.e. PLEASE ANSWER AS IF HEARING VOICES IS NEGATIVE AND SEEING THINGS IS POSITIVE or PLEASE ANSWER AS IF HEARING VOICES IS POSITIVE AND SEEING THINGS IS NEGATIVE). A full list of label stimuli, target stimuli, and response options for the IRAP is provided in Table 5.

Table 5

Stimuli and Response Options of the Evaluation IRAP employed in Experiment 3

Evaluation IRAP	
Hearing Voices is	Seeing Things is
Safe	Bad
Fine	Dangerous
Interesting	Wrong
True	False

Procedure

The current study comprised of two stages, one involving the implicit measure (the IRAP) and the second involving the explicit measures. Prior to the IRAP, non-voice hearing participants were identified using Item No. 33 of the CAPE and were subsequently provided with a written explanation of the phenomenon of voice hearing. This information included the following:

There are phenomena within human psychology called ‘hearing voices’ or ‘seeing things’, more formally known as auditory and visual hallucinations. They are often encompassed within the diagnosis of Schizophrenia or Psychosis, but often are present in those without any clinical diagnosis. When someone ‘hears things’ or ‘sees things’, these things are not really there, they are in fact an hallucination. It is these phenomena that the current research is investigating.

All participants completed the IRAP prior to the explicit measures. The verbal and automated instructions provided to participants for completing the IRAP

were identical to Experiment 1. For illustrative purposes, see Figure 10 for a schematic representation of the screen presentation of the Evaluation IRAP.

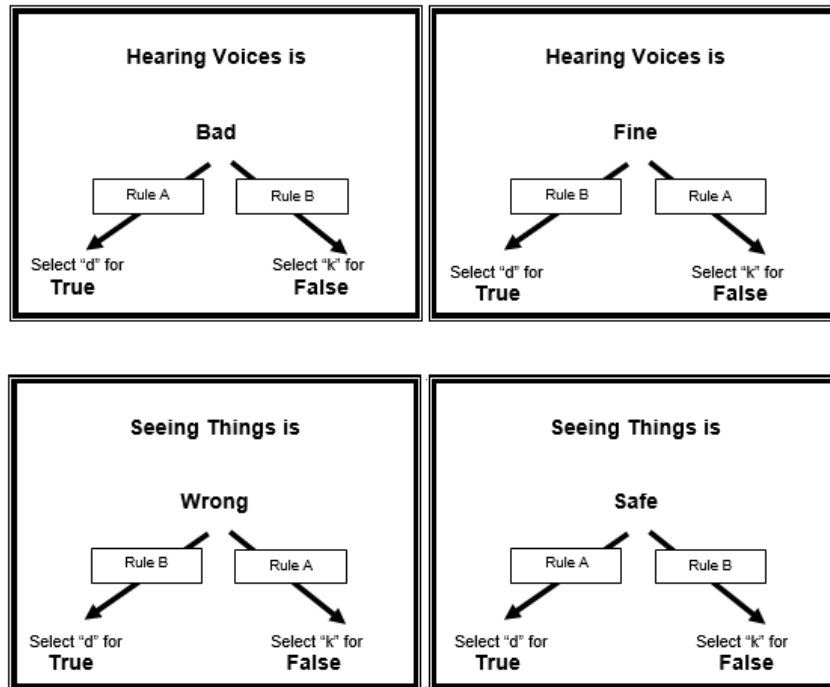


Figure 10. Schematic representation of the four trial-types presented in the Evaluation IRAP in Experiment 3. The arrows and text boxes did not appear on the participant’s screen, they indicate the correct responses for Rule A and Rule B blocks of trials. **The labels used for the four trial-types are as follows: Voices-Negative (top-left), Voices-Positive (top-right), Visions-Negative (bottom-left) and Visions-Positive (bottom-right).**

Once participants finished the IRAP, they completed the five explicit measures in a pre-determined sequence (AHRS, CAPE, AAQ-II, ATQ and SAB).

Ethical Considerations

The steps taken to circumvent any potential ethical concerns can be summarised as: any participant who reported a history of voice hearing (on the AHRS) was not exposed to the IRAP. All other ethical aspects of Experiment 3 were identical to Experiment 1.

Results

Analytic Strategy

Given that the primary aim of the current study was to assess reactions to voices, all data from the visions trial-types was excluded from analyses. This exclusion was also based on the fact that there were no explicit measures that assessed visions directly. To assess the potential effects of trial-type, a one-way ANOVA was conducted for the Evaluation IRAP. Exploratory analyses using the CAPE investigated the potential influence of psychotic-like experiences on IRAP effects. Data was split into two groups using a median split on the positive dimension subscale of the CAPE. The median split was conducted as only four participants fell over the recommended cut-off for being at-risk of psychosis (Mossaheb et al., 2012) . A mixed between within ANOVA was then conducted. Correlational analyses were then carried out to investigate the potential predictive validity of the IRAP in this context. Before conducting the IRAP analyses, summaries of the explicit data are provided in the section below.

Explicit Measures Data

The means and standard deviations were calculated from each participant's responses on each of the explicit measures and data are summarised in Table 6.

Table 6

Explicit Data Summary for Experiment 3

Scales	<i>Mean (SD)</i>
AAQ-II	23.20 (6.27)
CAPE	
CAPE positive dimension	2.65 (0.353)
CAPE depressive dimension	3.49 (0.99)
CAPE negative dimension	3.72 (0.94)
ATQ	
PMDC	11.00 (5.78)
NSNE	12.80 (5.85)
LSE	2.73 (1.58)
Helplessness	3.80 (1.78)
Believability	66.00 (23.99)
SAB	23.33 (3.87)

*Note. Maximum scores are: AAQ-II = 70; CAPE dimensions = 6; ATQ: PMDC = 25; NSNE = 35; LSE = 10; Helplessness = 10; Believability = 150; SAB = 56.

On the AAQ-II, means revealed that participants overall had low rates of inflexibility. On the CAPE, means revealed that participants overall had low to moderate rates of psychotic-like experiences. On the ATQ, means revealed low to moderate rates of automatic thoughts. And on the SAB, means revealed low stigmatising attitudes toward mental illness.

IRAP Data

Scoring of the IRAP was always conducted using the standardised approach for transforming latency data into D_{IRAP} scores (see Nicholson & Barnes-Holmes, 2012a). All data from any participant which fell below 80% accuracy and above 2000ms latency on any of the six test blocks were omitted from the dataset ($N=6$). The final dataset comprised of 29 participants, 11 were male and 19 were female.

The mean D_{IRAP} scores for the IRAP are presented in Figure 11 (visions trial-types are excluded). On the Voices-Positive trial-type, participants showed pro-voices effects. And a similar pattern emerged on the Voices-Negative trial-type, although this effect was only marginal.

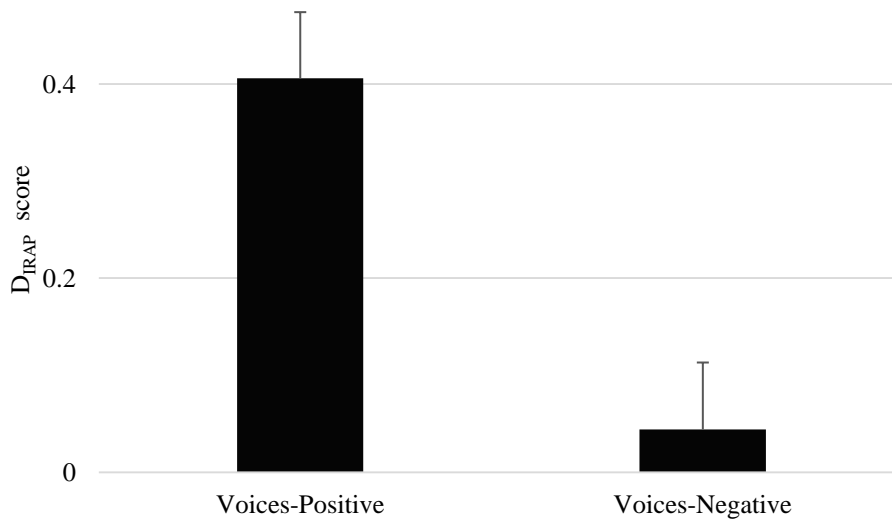


Figure 11. Mean D_{IRAP} scores for the three groups on the Evaluation IRAP in Experiment 3. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

In order to investigate the effects of group on trial-type, a one-way ANOVA was conducted and a main effect for trial-type was found ($df=1$, $F=16.471$, $p<0.001$, $\mu^2=0.985$). Post-hoc trial-type analyses (two one-sample t-tests) investigated whether the D_{IRAP} trial-type effects differed significantly from zero and Voices-Positive did ($df=28$, $t=5.961$, $p<0.0001$).

CAPE analyses. The mean D_{IRAP} scores for high and low CAPE are presented in Figure 12. For Voices-Positive, the low group showed greater pro-voices effects than the high group. For Voices-Negative, the low group showed marginal anti-voices effects, whereas the high group showed pro-voices effects.

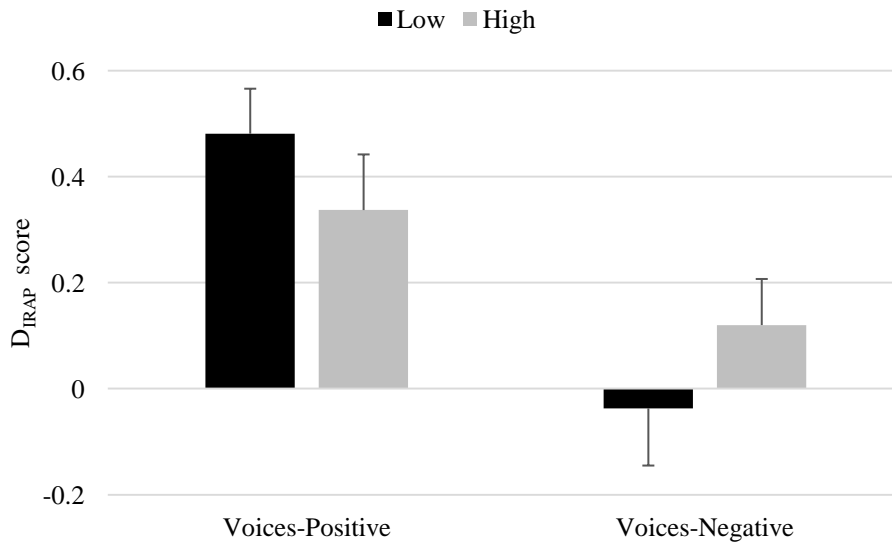


Figure 12. Mean D_{IRAP} scores for the high and low CAPE groups in the Evaluation IRAP in Experiment 3. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

A mixed between within ANOVA demonstrated a significant main effect for trial-type ($F=18.169$, $p<0.001$, $\mu^2=0.992$). Post-hoc analyses as two independent t-tests found no significant differences between the two CAPE groups (all $p's>0.05$). Again, four one-sample t-tests investigated whether the effects in each group differed significantly from zero. For both groups, Voices-Positive was significant (low group: $df=13$, $t=5.668$, $p<0.0001$; high group: $df=14$, $t=3.209$, $p<0.01$).

Correlations

A correlation matrix was conducted between the IRAP trial-types and the explicit measures. For Voices-Positive, a significant positive correlation was found with SAB stigma ($df=29$, $r=0.362$, $p<0.05$). That is, the greater the stigma, the greater the pro-voices effect. For Voices-Negative, significant positive correlations were recorded with: personal maladjustment and desire for change (PMDC, $df=29$, $r=0.455$, $p<0.01$); negative self-concepts and negative expectations (NSNE, $df=29$,

$r=0.438, p<0.05$); and helplessness ($df=29, r=0.447, p<0.01$). That is, the higher the scores on PMDC, NSNE and Helplessness, the greater the anti-voices effect.

Bonferroni adjustment. Due to multiple correlational analyses conducted, a Bonferroni adjustment of the alpha level was applied ($p<0.0005$). However, using this p-value, the above correlations do not reach significance (all $p's>0.0005$). Indeed, while this adjustment is deemed necessary for statistical conservatism and to avoid the occurrence of a Type 1 error, for the purposes of the current thesis, the correlations which were significant at the 0.05 level will also be discussed.

Summary and Discussion

Experiment 3 was the first to assess explicit and implicit attitudes toward hearing voices, using an Evaluation IRAP that juxtaposed hearing voices with seeing things using positive versus negative comparisons, and presented to participants. Overall, hearing voices was positively evaluated. The distress level analyses showed that participants high on psychotic-like experiences (on the CAPE) evaluated voices positively, while those who were low were even more positive, but also showed marginal negativity. Indeed, in this study, *all* effects showing positivity toward voices were significant. Surprisingly, implicit positivity toward voice hearing correlated with high stigma, but again negativity correlated with high distress. Overall, the findings appear to contrast findings reported by Peris and colleagues (2008), where there was stigma toward schizophrenia. However, the current sample comprised of a high proportion psychology undergraduates exposed to some level of psychological training, which Peris et al. had already shown to reduce stigma.

Although this was the first of our studies to target voice hearing directly, the findings were largely consistent with the previous two studies, and participants once

again showed unexpected implicit positivity towards unusual psychological experiences, such as voice hearing. That is, this is now the third study of this research programme which shows that individuals, with typically low psychological distress, do not show the implicit stigma toward psychological suffering that is typically assumed. But again, the data suggest a more complex picture in which an individual's explicit distress does appear to influence positivity or negativity to these psychological experiences. For example, once again, this study showed that implicit negativity to hearing voices predicted participants' distress.

Limitations

There are a number of limitations of Experiment 3 which should be reflected on to guide future research. 1. The target stimuli were selected to assess implicit positivity or negativity towards voice hearing, however, once again, this may not sufficiently capture stigmatisation toward voices. Future studies could include stimuli which target stigmatising behaviours toward voice hearers (e.g. I would stay away from these people, they are different to me etc.). 2. The current sample comprised of a high proportion of psychology undergraduates exposed to some level of psychological training, which may account for some of the positivity, therefore, it would be interesting to replicate this study in a sample with no training in psychology.

Chapter 4

Experiment 4

Assessing the effects of a voice hearing simulation on IRAP effects in a non-voice hearing sample

Experiment 4

Assessing the Effects of a Voice Hearing Simulation on IRAP effects

In recent years, the growing literature on voice hearing has begun to focus on the possible malleability of negative professional attitudes to psychological suffering. In the context of voice hearing, several studies have shown that simulation procedures can successfully reduce stigma and improve attitudes toward voices (Dearing & Steadman, 2009; Deegan, 1996; Kidd et al., 2015; Wieland et al., 2015; Wilson et al., 2009). However, several studies have shown that this positive outcome may depend on other variables, such as level of professional experience (Sideras et al., 2015). Nonetheless, some studies have reported negative post-simulation outcomes, in which participants show increased social distance from voice hearers (Brown et al., 2010; Kalyanaraman et al., 2010).

Experiment 4 sought to investigate potential changes in implicit stigmatisation or fear of hearing voices in participants subjected to a hearing voices simulation and a Fear IRAP. Given the mixed outcomes from these simulations as reported in the literature, and the general implicit positivity to voice hearing, as observed in Experiment 3, it was difficult to predict whether any change would occur in the positive IRAP effects from baseline to post-simulation, however, we hypothesised that we may observe a reduction in positivity at post-simulation.

Method

Setting

All aspects of the setting in Experiment 4 were identical to Experiment 1.

Participants

The current study involved a group of non-voice hearing participants who were identified as such using current screening methods from a general sample of undergraduate students recruited from MU. Forty-three non-voice hearing individuals were identified using current screening methods, 22 of these were male and 21 were female, with an age range of 18 to 28 years and a mean age of 19.72 years and a standard deviation of 1.81 years.

Materials

Explicit measures. Three broad categories of explicit measures were administered. The first series of measures assessed voice hearing (the AHRS) and psychosis (CAPE). The second set of measures more broadly assessed general psychological well-being (AAQ-II, ATQ and the DASS). The third measured stigma toward mental health difficulties (SAB). See Chapter 2 for outlines of the DASS and SAB, and see Chapter 3 for outlines of the AHRS, CAPE, AAQ-II and the ATQ.

The IRAP. The IRAP was administered on a standard desktop computer. This was used to present the instructions and stimuli, and to record responses. The current study involved one IRAP which assessed fearful or normative evaluations of voices (referred to as the Fear IRAP). Once again, the normative stimuli were intentionally selected to be more ambivalent than highly positive, as we hypothesised that this would be unlikely in non-voice hearing participants who have never had these experiences.

The IRAP contrasted hearing voices with seeing things, using the labels HEARING VOICES IS and SEEING THINGS IS. Each trial-type presented one of these two types of category labels. These were accompanied by one of three positive

(e.g. FINE) or three negative target stimuli (e.g. SCARY). The screen also presented two response options, TRUE and FALSE. Before each block of trials, the screen presented one of two rules for responding (i.e. PLEASE ANSWER AS IF HEARING VOICES IS SCARY AND SEEING THINGS IS OKAY or PLEASE ANSWER AS IF HEARING VOICES IS OKAY AND SEEING THINGS IS SCARY). A full list of label stimuli, target stimuli, and response options for the IRAP is provided in Table 7.

Table 7

Stimuli and Response Options of the Fear IRAP employed in Experiment 4

Fear IRAP	
Hearing Voices is	Seeing Things is
Okay	Scary
Fine	Distressing
Grand	Worrying
True	False

Voice hearing simulation. The voice hearing simulation comprised of two phases, the first phase comprised of critical voices commenting on behaviour (e.g. “You are so stupid”) and the second phase comprised of arbitrary abnormal sounds. Both phases of the simulation were produced by a voice hearer to simulate their own heard voices. The simulation lasted for 2 minutes in total, one minute for each phase of voices/sounds.

Procedure

The current study had a repeated measures design and comprised of four stages: 1. The IRAP; 2. An audio simulation of a voice hearing experience; 3. A second exposure to the IRAP and; 4. The explicit measures. The four stages are outlined in Figure 13.

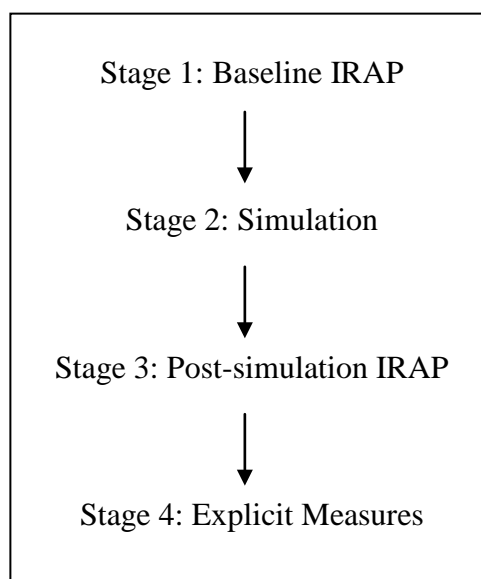


Figure 13. The four stages in Experiment 4.

Prior to the first exposure to the IRAP, non-voice hearing participants were identified using Item No. 33 of the CAPE and were subsequently provided with a written explanation of the phenomenon of voice hearing. See Chapter 3 for an example of this explanation.

Stage 1: Baseline IRAP. All participants completed the Fear IRAP for the first time in Stage 1. The verbal and automated instructions provided to participants for completing the IRAP were identical to Experiment 1. For illustrative purposes, see Figure 14 for a schematic representation of the screen presentation of the IRAP.

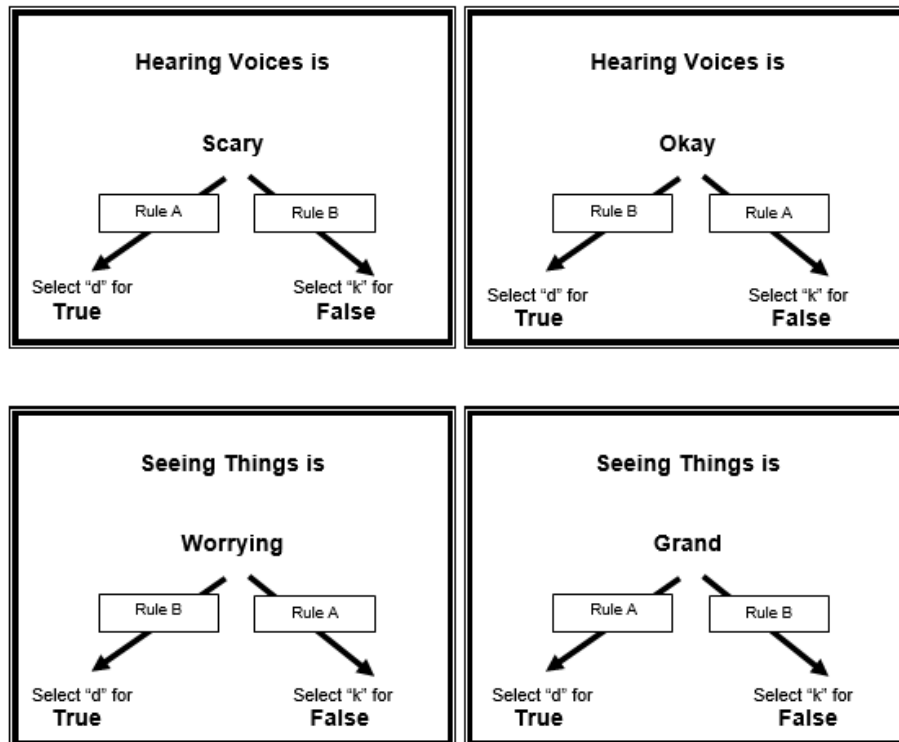


Figure 14. Schematic representation of the four trial-types presented in the Fear IRAP in Experiment 4. The arrows and text boxes did not appear on the participant’s screen, they indicate the correct responses for Rule A and Rule B blocks of trials. **The labels used for the four trial-types are as follows: Voices-Scary (top-left), Voices-Okay (top-right), Visions-Scary (bottom-left) and Visions-Okay (bottom-right).**

Stage 2: Voice hearing simulation. Once participants had completed the IRAP, the voice hearing simulation was presented. This involved listening to an mp3 file through headphones at a pre-experimentally set volume.

Stage 3: Post-simulation IRAP. After the voice hearing simulation, participants completed the second exposure to the Fear IRAP.

Stage 4: Explicit measures. Once participants finished the second exposure to the IRAP, they completed the six explicit measures in a pre-determined sequence (AHRs, CAPE, AAQ-II, ATQ, DASS and SAB).

Ethical Considerations

This experiment included the addition of a voice hearing simulation and the steps taken to circumvent any potential ethical concerns can be summarised as: any participant who reported a history of voice hearing (on the AHRS) was not exposed to the IRAP or the voice hearing simulation. All other aspects of potential ethical considerations in Experiment 4 were identical to Experiment 3.

Results

Analytic Strategy

Given that the primary aim of the current study was to assess reactions to voices, all data from the visions trial-types were again excluded from the analyses. Hence, mixed between within ANOVAs were conducted for each IRAP exposure (pre and post) and a repeated measures ANOVA investigated potential effects of simulation on IRAP scores. Exploratory analyses using the CAPE investigated the potential influence of psychotic-like experiences on IRAP effects. Data was split into two groups using a median split on the positive dimension subscale of the CAPE. The median split was conducted as only three participants fell over the recommended cut-off for being at-risk of psychosis (Mossaheb et al., 2012) . This analysis involved a mixed factorial ANOVA. Correlational analyses investigated the potential predictive validity of the IRAP in this context.

Before conducting the IRAP analyses, summaries of the explicit data are provided in the section below.

Explicit Measures Data

The means and standard deviations were calculated from each participant's responses on each of the explicit measures and data are summarised in Table 8.

Table 8

Explicit Data Summary for Experiment 4

Scales	Mean (SD)
AAQ-II	21.43 (9.10)
CAPE	
CAPE positive dimension	1.94 (0.806)
CAPE depressive dimension	2.85 (1.05)
CAPE negative dimension	3.09 (1.33)
DASS	12.607 (8.40)
Depression	3.79 (3.41)
Anxiety	3.29 (3.11)
Stress	5.54 (3.46)
SAB	23.33 (3.87)

*Note. Maximum scores are: AAQ-II = 70; CAPE dimensions = 6; DASS total = 126; DASS subscales = 42; SAB = 56.

On the AAQ-II, means revealed that participants overall had low rates of inflexibility. On the CAPE, means revealed that participants overall had low to moderate rates of psychotic-like experiences. For the DASS, means revealed that participants overall had low rates of depression, anxiety and stress. And on the SAB, means revealed low stigmatising attitudes toward mental illness.

IRAP Data

Scoring of the IRAP was always conducted using the standardised approach for transforming latency data into D_{IRAP} scores (see Nicholson & Barnes-Holmes,

2012a). All data from any participant which fell below 80% accuracy and above 2000ms latency on any of the six test blocks in each IRAP were omitted from the dataset ($N=15$). The final dataset comprised of 28 participants.

The mean D_{IRAP} scores on the *Baseline IRAP* are presented in Figure 15 (visions trial-types are excluded). On the Voices-Okay trial-type, participants showed pro-voices effects, and on the Voices-Scary trial-type, participants showed marginal anti-voices effects.

The mean D_{IRAP} scores on the *Post-Simulation IRAP* are also presented in Figure 15. Similar to the baseline IRAP, on Voices-Okay, participants showed pro-voices effects and on Voices-Scary, participants showed anti-voices effects.

Figure 15 indicates that the simulation was associated with changes in both trial-types from baseline to post-simulation. From baseline to post-simulation, the pro-voices effect on Voices-Okay decreased, while the anti-voices effect on Voices-Scary increased.

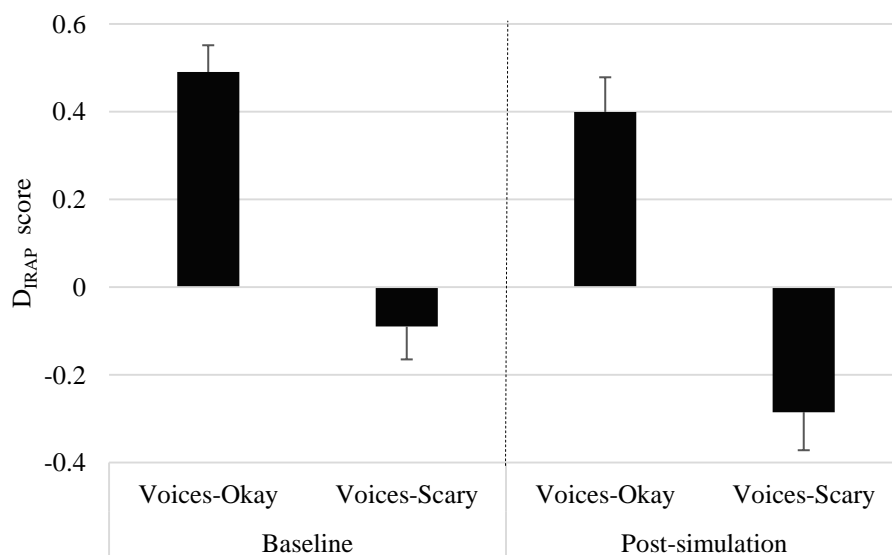


Figure 15. Mean D_{IRAP} scores at baseline and post-simulation on the Fear IRAP in Experiment 4. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

In order to investigate the effects of trial-type at *Baseline*, a one-way ANOVA was conducted and a main effect was found ($df=27$, $F=77.771$, $p<0.0001$, $\mu^2=1.0$). Post-hoc analyses (two one-sample t-tests) indicated that Voices-Okay was significant ($df=27$, $t=-8.040$, $p<0.0001$).

At *Post-simulation*, a one-way ANOVA again found a main effect ($df=27$, $F=48.395$, $p<0.0001$, $\mu^2=1.0$). Again, two one-sample t-tests indicated that *both* Voices-Okay ($df=27$, $t=-5.066$, $p<0.0001$) and Voices-Scary were significantly different from zero ($df=27$, $t=3.291$, $p<0.01$).

Repeated measures analyses. To investigate the effect of simulation on trial-type, a repeated measures ANOVA was conducted, and found a main effect for trial-type ($df=1$, $F=114.183$, $p<0.0001$, $\mu^2=1.0$). However, post-hoc analyses as two dependent t-tests showed no differences from baseline to post-simulation (all $p's>0.05$).

CAPE analyses. The mean D_{IRAP} scores for high and low positive dimension CAPE scores on the IRAP are presented in Figure 16. On the baseline IRAP, the high group showed greater pro-voices effects than the low group on Voices-Okay. For Voices-Scary, the low group showed anti-voices effects, whereas the high group showed marginal pro-voices effects. On the post-simulation IRAP, the low group showed greater pro-voices effects than the high group on Voices-Okay. And for Voices-Scary, the high group showed greater anti-voices effects than the low group. In summary, the IRAP effect for participants with low psychotic-like experiences remained the same from baseline to post-simulation, but the IRAP effects for participants with high psychotic-like experiences became less pro-voices on Voices-Okay and more anti-voices on Voices-Scary.

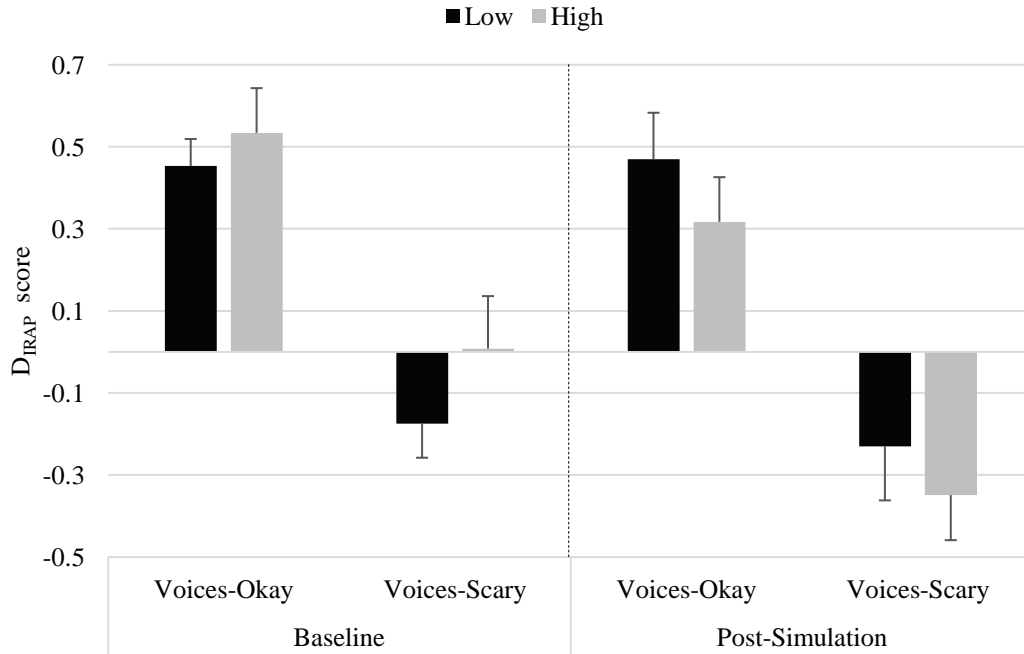


Figure 16. Mean D_{IRAP} scores at baseline and post-simulation on the Fear IRAP for high and low CAPE groups scores in Experiment 4. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

A mixed factorial ANOVA demonstrated a significant main effect for trial-type ($F=33.885$, $p<0.0001$, $\eta^2=1.0$). Post-hoc analyses as four independent t-tests found no significant differences between the two CAPE groups (all p 's>0.05). Four dependent t-tests investigated potential differences between baseline and post-simulation IRAP effects for each group. The only differences were found in the high group whose effects on Voices-Scary differed significantly between baseline and post-simulation ($df=12$, $t=-2.190$, $p<0.05$). Again, eight one-sample t-tests investigated whether the effects in each group differed significantly from zero. For the low group, significant effects were found for baseline Voices-Okay ($df=12$, $t=-6.849$, $p<0.0001$) and Voices-Scary ($df=12$, $t=2.113$, $p<0.05$), and for post-simulation Voices-Okay ($df=12$, $t=-4.161$, $p<0.01$). For the high group, significant effects were found for baseline Voices-Okay ($df=12$, $t=-4.903$, $p<0.001$), and for

post-simulation Voices-Okay ($df=12$, $t=-2.915$, $p<0.05$) and Voices-Scary ($df=12$, $t=3.167$, $p<.01$).

Correlations

A correlation matrix was conducted between the IRAP baseline and post-simulation trial-types and the explicit measures. No significant correlations were found between the trial-types and the explicit data (all $p's>0.05$). When split by CAPE, for the high group, there was a significant negative correlation between post-simulation Voices-Okay and the CAPE positive dimension ($df=13$, $r=-0.586$, $p<0.05$). That is, the higher the psychotic-like experiences, the lower the pro-voices effects.

Bonferroni adjustment. Due to multiple correlational analyses conducted, a Bonferroni adjustment of the alpha level was applied ($p<0.0005$). However, using this p-value, the above correlations do not reach significance (all $p's>0.0005$). Indeed, while this adjustment is deemed necessary for statistical conservatism and to avoid the occurrence of a Type 1 error, for the purposes of the current thesis, the correlations which were significant at the 0.05 level will also be discussed.

Summary and Discussion

This experiment sought to investigate the malleability of fearful evaluations of hearing voices, using a Fear IRAP and a voice hearing simulation. At baseline and post-simulation, hearing voices was implicitly evaluated as both positive *and* fearful, although positivity toward voices reduced and fear increased after the simulation, and this latter fearful effect was significant.

Once again, the IRAP indicated implicit positivity by participants to voice hearing, tinged with elements of implicit fear, hence less negativity or stigma than one might expect. But again, this appears to be influenced by participants' own experiences. The simulation procedure employed in Experiment 4 appeared to be particularly useful at shedding light between the complex relationship between this implicit positivity and participants own experiences in terms of assessing the malleability of the IRAP outcomes. Indeed, it was particularly interesting that the implicit positivity of participants who had low levels of their own psychotic-like experiences changed little as a result of directly experiencing voices through the simulation, while those perhaps more familiar with psychotic-like experiences became more fearful and less positive. In this study, the implicit positivity in both groups at both time points was significant, and while the implicit negativity in the low group was significant at baseline, it was the negativity in the high group that was significant at post-simulation. Overall, these findings appear to contrast the positive effects found in the literature where there were reductions in stigma after a simulation (Dearing & Steadman, 2009; Deegan, 1996; Kidd et al., 2015; Sideras et al., 2015; Wieland, et al., 2015; Wilson et al., 2009), but appear complement those studies which have found negative simulation outcomes (e.g. Brown et al., 2010; Kalyanaraman et al., 2010).

The four studies conducted in the thesis thus far provided a useful benchmark for understanding the implicit attitudes of participants to various patterns of psychological suffering or related concepts. This benchmark allowed us to ascertain whether effects that might be recorded with individuals with direct suffering pertained directly to that suffering or whether these are attitudes held by the general population. Having now established this baseline of attitudes, which was indeed

unexpectedly positive in each study, we naturally turned our attention toward determining whether similar or different effects would be recorded with voice hearers. This was the focus of the four remaining experiments in the thesis.

Limitations

There are a number of limitations of Experiment 4 which should be reflected on to guide future research. 1. While assessing stigma was not the primary aim of this study, in order to be able to compare the impact of a simulation procedure on implicit stigma against existing studies more concisely, future studies could include stimuli which target stigmatising behaviours toward sufferers (e.g. I would stay away from these people, they are different to me etc.). 2. The current sample comprised of a high proportion of psychology undergraduates exposed to some level of psychological training, which may account for some of the positivity, therefore, it would be interesting to replicate this study in a sample with no training in psychology.

Chapter 5
Experiment 5

Assessing implicit perceptions of
the normality of voice hearing in
non-clinical voice hearers and non-
voice hearing controls

Experiment 5

Assessing Implicit Perceptions of the Normality of Voice hearing

Various studies suggest that mental health labels (as an indication of ‘abnormality’) actually facilitate stigma, especially when experiences (such as voice hearing) are perceived as rare or unusual (Corrigan, 2004; Mak, Poon, Pun, & Cheung, 2007). However, no studies have focused specifically on the perceived normality of voices per se, either in the context of stigma or self-stigma in sufferers. Experiment 5 sought to investigate implicit evaluations of non-clinical voice hearers’ using the Normality IRAP. Our key aim was to compare this group against non-voice hearing control participants (non-voice hearers). Based on the previous IRAPs, we hypothesised some implicit positivity by control participants, but given that this was the first sample of voice hearers, it was hard to predict whether or not this latter group would show different IRAP effects relative to the controls, however, we hypothesised less implicit positivity by voice hearers.

Method

Setting

All aspects of the setting in Experiment 5 were identical to Experiment 1.

Participants

The current study involved two groups of participants from a general sample of undergraduate students recruited from MU. One group was categorised as non-clinical voice hearers and the other group comprised of a non-voice hearing control group. Seven non-clinical voice hearers were identified using current screening

methods, four of these were male and 3 were female. Thirty-six non-voice hearing individuals were identified, 11 of these were male and 25 were female. In total, the study involved 43 participants, 15 males and 21 females, with an age range of 18 to 38 years and a mean age of 22.16 years and a standard deviation of 2.76 years.

Materials

Explicit measures. Three broad categories of explicit measures were administered. The first series of measures assessed voice hearing (AHRS, the Beliefs about Voices Questionnaire, and the Voices Acceptance and Action Questionnaire) and psychosis (CAPE). The second set of measures more broadly assessed general psychological well-being (AAQ-II and ATQ). The third measured stigma toward mental health difficulties (SAB). See Chapter 2 for full outlines of the SAB and see Chapter 3 for full outlines of AHRS, CAPE, AAQ-II and the ATQ.

Beliefs About Voices Questionnaire–Revised (BAVQ-R; Chadwick, Lees, & Birchwood, 2000). The BAVQ-R is a 35-item scale that targets beliefs, feelings and behaviours about voice hearing (e.g. “my voice is punishing me for something I have done”). The measure comprises seven subscales: malevolence; benevolence; omnipotence; emotional resistance; behavioural resistance; emotional engagement; and behavioural engagement. All items are rated on a 4-point scale from 0 (disagree) to 3 (strongly agree). Subscales are scored independently and indicate high or low levels of each dimension. The BAVQ-R yields a minimum score of 0 for all subscales, and a maximum score of: 18 for malevolence, benevolence and omnipotence; 12 for emotional resistance, emotional engagement and behavioural engagement; and 15 for behavioural resistance. The BAVQ-R subscales have

demonstrated adequate internal consistency with an alpha coefficient of = 0.74 to 0.88 (Chadwick et al.).

Voices Acceptance and Action Questionnaire (VAAS; Shawyer et al., 2007). The VAAS is a 31-item scale that measures acceptance of voices (e.g. “I accept the fact that I hear voices”). The scale comprises two broad sections that measure emotional acceptance and behavioural acceptance. All items are rated on a 5-point scale from 1 (strongly disagree) to 5 (strongly agree). The VAAS yields an overall score with a maximum of 155 indicating high voice acceptance and a minimum of 0 indicating low acceptance. This scale has demonstrated excellent internal consistency with an alpha coefficient of 0.90 (Shawyer et al.).

The IRAP. The IRAP was administered on a standard desktop computer. This was used to present the instructions and stimuli, and to record responses. The current study involved one IRAP which assessed normality or abnormality of voices, hereafter referred to as the Normality IRAP.

The IRAP contrasted hearing voices with seeing things, using the labels HEARING VOICES IS and SEEING THINGS IS. Each trial-type presented one of these two types of category labels. These were accompanied by one of three positive (e.g. NORMAL) or three negative target stimuli (e.g. ABNORMAL). The screen also presented two response options, TRUE and FALSE. Before each block of trials, the screen presented one of two rules for responding (i.e. PLEASE ANSWER AS IF HEARING VOICES IS ABNORMAL AND SEEING THINGS IS NORMAL or PLEASE ANSWER AS IF HEARING VOICES IS NORMAL AND SEEING THINGS IS ABNORMAL). A full list of label stimuli, target stimuli, and response options for the IRAP is provided in Table 9.

Table 9

Stimuli and Response Options of the Normality IRAP employed in Experiment 5

Normality IRAP	
Hearing Voices is	Seeing Things is
Normal	Abnormal
Sane	Insane
Acceptable	Crazy
True	False

Procedure

The current study comprised of two stages, one involving the implicit measure (the IRAP) and the second involving the explicit measures. Thus, the study had a between groups repeated measures design. Prior to the IRAP, non-clinical voice hearing and non-voice hearing participants were identified using Item No. 33 of the CAPE. Non-voice hearing participants were subsequently provided with a written explanation of the phenomenon of voice hearing (see Chapter 3).

All participants completed the IRAP prior to the explicit measures. The verbal and automated instructions provided to participants for completing the IRAP were identical to Experiment 1. For illustrative purposes, see Figure 17 for a schematic representation of the screen presentation of the Normality IRAP.

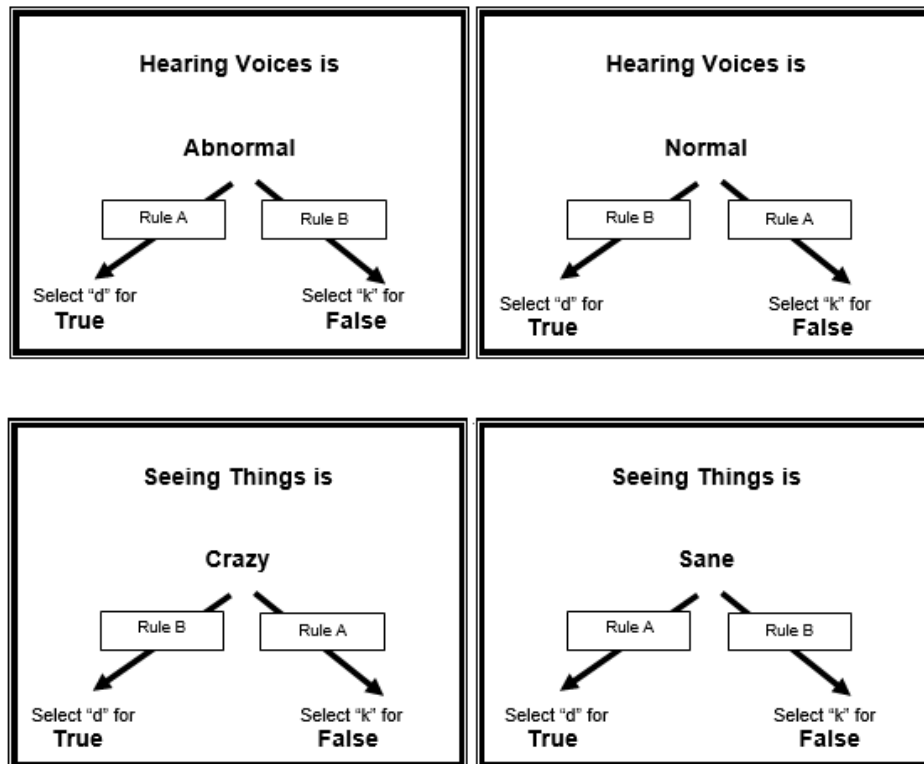


Figure 17. Schematic representation of the four trial-types presented in the Normality IRAP in Experiment 5. The arrows and text boxes did not appear on the participant's screen, they indicate the correct responses for Rule A and Rule B blocks of trials. **The labels used for the four trial-types are as follows: Voices-Abnormal (top-left), Voices-Normal (top-right), Visions-Abnormal (bottom-left) and Visions-Normal (bottom-right).**

Once participants finished the IRAP, the voice hearing participants completed the seven explicit measures in a pre-determined sequence (AHRS, BAVQ-R, VAAS, CAPE, AAQ-II, ATQ and SAB), while the non-voice hearing participants completed the non-voice related measures in a pre-determined sequence (CAPE, AAQ-II, ATQ and SAB).

Ethical Considerations

This experiment included the addition of a voice hearing sample and the steps taken to circumvent any potential ethical concerns can be summarised as follows. Identical to the control participants, any voice hearing participant who reported a

history of psychological distress were not exposed to the IRAP. All other aspects of potential ethical considerations in Experiment 5 were identical to Experiment 3.

Results

Analytic Strategy

As in the previous chapters, all data from the visions trial-types were excluded from the analyses. To assess potential differences in the IRAP effects between the two groups of participants (i.e. non-clinical voice hearers and non-voice hearing controls), a mixed between within ANOVA was conducted. Due to the small sample size of non-clinical voice hearers employed in the study, exploratory analyses using the CAPE investigated the potential influence of psychotic-like experiences on IRAP effects. Data from the non-voice hearers was split into two groups using a median split on the positive dimension subscale of the CAPE and compared to voice hearers. The median split was conducted as only four participants fell over the recommended cut-off for being at-risk of psychosis (Mossaheb et al., 2012). This was then involved in a mixed between within ANOVA. Correlational analyses were used to investigate the potential predictive validity of the IRAP in this context.

Before conducting the IRAP analyses, summaries of the explicit data are provided in the section below.

Explicit Measures Data

The means and standard deviations were calculated from each participant's responses on each of the explicit measures and data are summarised in Table 10.

Table 10

Explicit Data Summary for Experiment 5

Scales	Voice Hearers	Controls	T-Test
	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>T</i>
AAQ-II	24.57 (10.01)	22.66 (8.97)	.497
CAPE			
CAPE positive dimension	2.45 (0.48)	2.0 (0.61)	1.837
CAPE depressive dimension	3.80 (1.58)	3.47 (1.18)	.627
CAPE negative dimension	3.58 (0.91)	3.13 (0.94)	1.149
BAVQ-R			
Malevolence	7.29 (4.42)	-	-
Benevolence	10.29 (6.75)	-	-
Omnipotence	10.00 (7.66)	-	-
Emotional Resistance	6.43 (4.35)	-	-
Behavioural Resistance	13.28 (8.01)	-	-
Emotional Engagement	7.29 (6.02)	-	-
Behavioural Engagement	6.43 (4.11)	-	-
VAAS	47.71 (5.91)	-	-
AHRS	9.57 (8.02)	-	-
ATQ			
PMDC	13.71 (5.88)	11.83 (4.29)	.973
NSNE	13.86 (5.52)	14.52 (5.57)	-.282
LSE	4.0 (2.31)	3.45 (2.10)	.613
Helplessness	4.0 (1.41)	3.69 (1.47)	.506
Believability	79.57 (32.33)	70.86 (27.18)	.735
SAB	24.43 (7.34)	23.38 (8.42)	.303

*Note. Maximum scores are: AAQ-II = 70; CAPE dimensions = 6; BAVQ-R: Mal, Ben & Omni = 18; Beh. Res. = 15, Emo. Res., Beh. Eng., Emo. Eng. = 12; VAAS = 155; AHRS = 44; ATQ: PMDC = 25; NSNE = 35; LSE = 10; Helplessness = 10; Believability = 150; SAB = 56. Missing values for explicit measures which were not administered to the control participants are denoted by “-”. Significance indicated by *($p < 0.05$).

For both groups, AAQ-II means revealed that participants overall had low rates of inflexibility. On the CAPE, means revealed that participants overall had low to moderate rates of psychotic-like experiences. For the ATQ, means revealed that participants overall had low to moderate rates of automatic thoughts. And on the SAB, means revealed low stigmatising attitudes toward mental illness.

For the voice hearers, on the AHRs, means revealed low voice severity. BAVQ-R means revealed moderate rates of malevolence, benevolence, omnipotence, emotional resistance, emotional engagement and behavioural engagement, and high rates on behavioural resistance. On the VAAS, means revealed low overall voice acceptance. Independent t-tests were conducted to investigate potential differences between the two groups, but none were found (all $p's > 0.05$).

IRAP Data

Scoring of the IRAP was always conducted using the standardised approach for transforming latency data into D_{IRAP} scores (see Nicholson & Barnes-Holmes, 2012a). All data from any participant which fell below 80% accuracy and above 2000ms latency on any of the six test blocks were omitted from the dataset ($N=7$). The final dataset comprised of 36 participants: 29 non-voice hearers; and seven non-clinical voice hearers.

Between groups analyses (non-clinical voice hearers and non-voice hearers). The mean D_{IRAP} scores for the IRAP are presented in Figure 18 (visions trial-types are excluded). On the Voices-Normal trial-type, both groups showed pro-voices effects. And a similar pattern emerged on the Voices-Abnormal trial-type. However, on both trial-types, non-clinical voice hearers had the greater pro-voices effect.

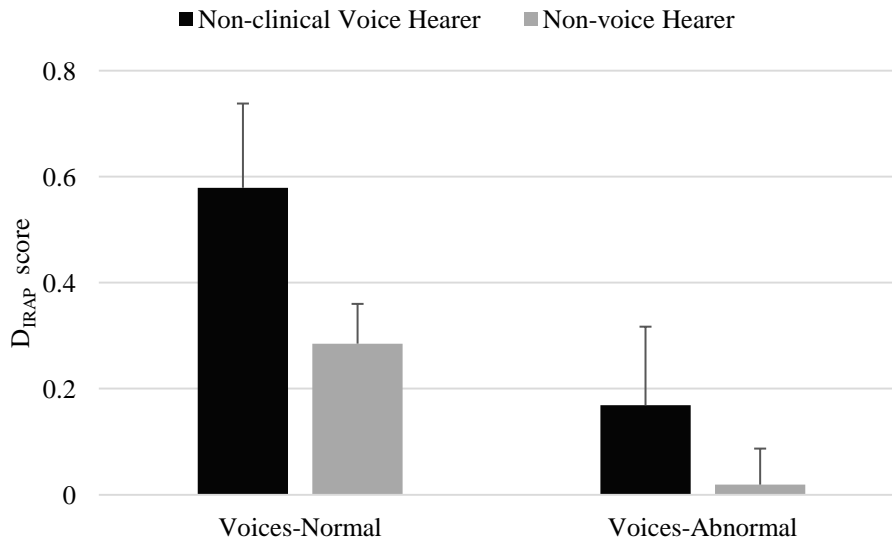


Figure 18. Mean D_{IRAP} scores for the two groups on the Normal IRAP in Experiment 5. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

In order to investigate the effects of group on trial-type, a mixed between within ANOVA found a main effect for trial-type ($df=1$, $F=8.935$, $p<0.01$, $\mu^2=0.842$). Post-hoc analyses, as two independent t-tests found no significance between the two groups on each trial-type (all $p's>0.05$). Furthermore, four one-sample t-tests investigated whether the D_{IRAP} trial-type effects differed significantly from zero and found that for both groups, Voices-Normal was significant (voice hearers: $df=6$, $t=3.639$, $p<0.01$; non-voice hearers: $df=28$, $t=3.811$, $p<0.001$).

CAPE analyses. For the CAPE analyses, data from the non-voice hearers was split into two groups using a median split on the positive dimension subscale, thus creating high vs. low CAPE non-voice hearing comparison groups. These were then also compared with the voice hearers, all of whom scored high on the CAPE. The mean D_{IRAP} scores for voice hearers ($N=7$) and high ($N=16$) and low ($N=13$) positive dimension CAPE scores (non-voice hearers) on the IRAP are presented in Figure 19. For Voices-Normal, the voice hearers showed the greatest pro-voices

effects, followed by the high CAPE group and then the low CAPE group. For Voices-Abnormal, the voice hearers again showed pro-voices effects (although modest), followed by the high CAPE group. The low CAPE group showed marginal anti-voices effects.

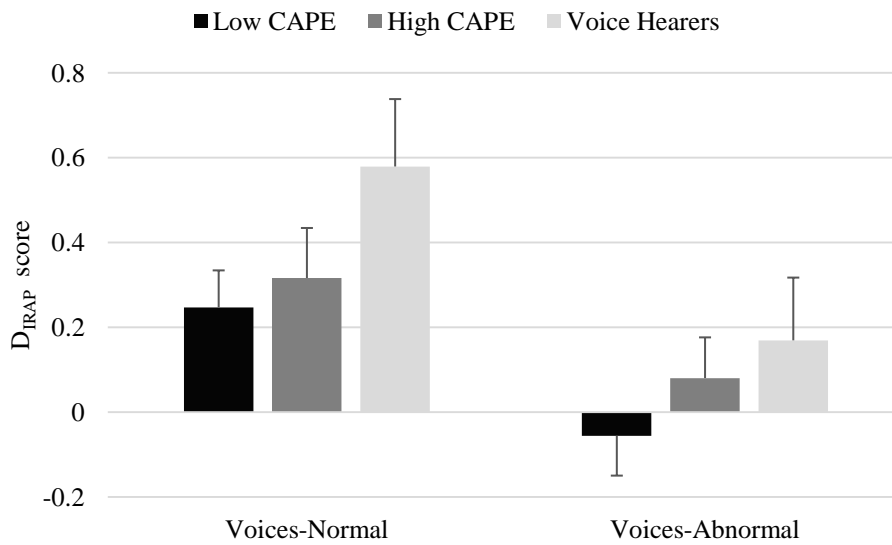


Figure 19. Mean D_{IRAP} scores on the Normal IRAP for voice hearers and high and low CAPE groups scores in Experiment 5. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

A mixed between within ANOVA (with group as the factor and trial-type as the measure) showed a significant main effect for trial-type ($F=10.774$, $p<0.01$, $\mu^2=0.907$). Post-hoc analyses, as six independent t-tests, investigated potential differences across groups on each trial-type, but none were found (all $p's>0.05$). Again, four one-sample t-tests (analyses completed in previous section for voice hearers) investigated whether the effects in each group differed significantly from zero. For both groups, significant effects were found for Voices-Normal (low: $df=12$, $t=2.829$, $p<0.01$; high: $df=15$, $t=2.688$, $p<0.01$).

Correlations

A correlation matrix was conducted between the IRAP and the explicit measures using the voice hearers' data. For Voices-Normal, a significant positive correlation was found with behavioural engagement on the BAVQ-R ($df=36$, $r=0.336$, $p<0.05$) and voice acceptance ($df=36$, $r=0.319$, $p<0.05$). That is, the greater the behavioural engagement with voices and acceptance of voices, the greater the pro-voices effect.

For Voices-Abnormal, a significant positive correlation was found with BAVQ-R benevolence appraisals ($df=36$, $r=0.331$, $p<0.05$) and emotional engagement with voices ($df=36$, $r=0.356$, $p<0.05$), with CAPE positive distress ($df=36$, $r=0.332$, $p<0.05$), the positive dimension ($df=36$, $r=0.333$, $p<0.05$), the depressive frequency ($df=36$, $r=0.331$, $p<0.05$), negative distress ($df=36$, $r=0.352$, $p<0.05$), negative frequency ($df=36$, $r=0.452$, $p<0.01$), and CAPE negative dimension ($df=36$, $r=0.408$, $p<0.01$). That is, the greater anti-voices effects on Voices-Abnormal, the greater: benevolence appraisals, emotional engagement with voices, positive, negative and depressive psychotic-like experiences.

Bonferroni adjustment. Due to multiple correlational analyses conducted, a Bonferroni adjustment of the alpha level was applied ($p<0.0005$). However, using this p-value, the above correlations do not reach significance (all p 's >0.0005). Indeed, while this adjustment is deemed necessary for statistical conservatism and to avoid the occurrence of a Type 1 error, for the purposes of the current thesis, the correlations which were significant at the 0.05 level will also be discussed.

Summary and Discussion

Experiment 5 was the first to involve non-clinical voice hearers and used the Normality IRAP to assess evaluations of hearing voices as normal. Both non-clinical voice hearers and non-voice hearing controls implicitly evaluated voice hearing as normal, and most importantly this effect was stronger for the voice hearers. This contrasted previous evidence that suffered categorised voice hearing as abnormal (Corrigan, 2004; Mak et al., 2007). Control participants who were high in psychotic-like experiences evaluated voices as most normal, while the low evaluated voices as least normal and showed marginal anti-normality reactions to voice hearing. Indeed, the normality effects found were significant for *all* groups in each analysis. Evaluations of voice hearing as normal correlated with behavioural engagement with voices as well as voice acceptance, while evaluations of voice hearing as abnormal correlated with high voice benevolence and high emotional engagement with voices and other psychotic-like experiences, which is largely consistent with the literature (Chadwick & Birchwood, 1994).

In support of our predictions and based on the four previous studies, Experiment 5 indicated once again that control participants had positive evaluations of voice hearing, at least in terms of categorising these experiences as normal (and not abnormal). Although it was initially difficult to predict whether there would be any differences between control participants and voice hearers, it was interesting to find that voice hearers more readily normalised this experience, even relative to the controls. Again, the distress level analyses allowed us parse to out these implicit attitudes and interestingly indicated that these normalisation effects were stronger for voice hearers and control participants who had high levels of psychotic-like experiences. The use of the two samples also allowed greater insight into other

factors predicted by the IRAP. For example, the fact that the implicit normalising correlated with acceptance of voices, while the categorisations of voices as abnormal correlated with voice engagement is very consistent with the literature on voice hearing. As a result, it seems fair to say that the utility of the IRAP is strongly supported in its ability to explore implicit attitudes to voice hearing and how this might distinguish voice hearers from controls.

Limitations

There are a number of limitations of Experiment 5 which should be reflected on to guide future research. 1. The time-point at which participants were hearing voices (i.e. past/present) was not controlled for and may have influenced the distress-level analyses. 2. The current sample comprised of a high proportion of psychology undergraduates exposed to some level of psychological training, which may account for some of the normality effects, therefore, it would be interesting to replicate this study in a sample with no training in psychology.

Chapter 6

Experiment 6

Assessing the acceptance of self and others hearing voices in non-clinical voice hearers and non-voice hearing controls

Experiment 6

Assessing Implicit Evaluations of Self and Others Hearing Voices

Corrigan (2004) proposed that the impact of self-evaluations must be considered when investigating appraisals of hearing voices among voice hearers, in terms of stigma against the “self” hearing voices *and* against “others” who hear voices. This would also be supported by the growing body of RFT literature, suggesting that self-evaluations play a key role in psychological suffering (e.g. Foody, Barnes-Holmes, & Barnes-Holmes, 2012). For RFT, the perspective-taking relations constitute the locus of control from which an individual views the self, others and the world. And it is these relations that facilitate a deeper understanding of the self and, in turn, psychological suffering. For example, psychological content (i.e. thoughts, feelings, memories etc.) can become problematic when it becomes *attached* to the self, where the self then becomes defined by this content.

To investigate and compare the potential stigma of non-voice hearing control participants and voice hearers toward voices as heard by the self and others, Experiment 6 presented these two groups with a Self IRAP and an Others IRAP. Again, our previous findings led us to predict that both groups would be largely positive toward voice hearing and Experiment 5 would suggest that greater positivity may be observed with the voice hearers. Given that we had not previously compared a self-versus-others IRAP, it was difficult to hypothesise whether there would be any differences between these two IRAPs and/or whether the two groups might differ in this regard.

Method

Setting

All aspects of the setting in Experiment 6 were identical to Experiment 1.

Participants

Recruited from a general sample of undergraduate students at MU, the current study involved two groups of participants. One group were categorised as non-clinical voice hearers and the other group comprised of a non-voice hearing control group. Twenty-four non-clinical voice hearers were identified using current screening methods. Fourteen of these were male and 10 were female. Forty-three non-voice hearing individuals were identified, 22 of these were male and 21 were female. In total, the study involved 67 participants, 36 males and 31 females, with an age range of 19 to 38 years and a mean age of 23.8 years and a standard deviation of 2.76 years.

Materials

Explicit measures. Three broad categories of explicit measures were administered. The first series of measures assessed voice hearing (AHRS, the BAVQ-R and the VAAS) and psychosis (CAPE). The second set of measures more broadly assessed general psychological well-being (AAQ-II and ATQ). The third measured stigma toward mental health difficulties (SAB). See Chapters 2, 3 and 5 for outlines of the explicit measures.

The IRAP. The IRAP was administered on a standard desktop computer. This was used to present the instructions and stimuli, and to record responses. The current study involved two IRAPs, one that assessed implicit appraisals regarding

the *self* hearing voices (referred to as the Self IRAP) and another that assessed appraisals regarding *other people* hearing voices (referred to as the Others IRAP).

The Self IRAP contrasted hearing voices with seeing things, using the labels IF I HEARD VOICES and IF I SAW THINGS. Each trial-type presented one of these two category labels, accompanied by one of three positive (e.g. IT WOULD BE FINE) or three negative target stimuli (e.g. IT WOULD BE SCARY). The screen also presented the response options TRUE and FALSE. Before each block of trials, the screen presented one of two rules for responding (i.e. PLEASE ANSWER AS IF HEARING VOICES IS OKAY AND SEEING THINGS IS SCARY or PLEASE ANSWER AS IF HEARING VOICES IS SCARY AND SEEING THINGS IS OKAY).

The Others IRAP contrasted hearing voices with seeing things, using the labels IF OTHER PEOPLE HEARD VOICES and IF OTHER PEOPLE SAW THINGS. Each trial-type presented one of these two category labels, accompanied by one of three positive (e.g. IT WOULD BE FINE) or three negative target stimuli (e.g. IT WOULD BE SCARY). The screen also presented the response options TRUE and FALSE. Before each block of trials, the screen presented one of two rules for responding (i.e. PLEASE ANSWER AS IF HEARING VOICES IS OKAY AND SEEING THINGS IS SCARY or PLEASE ANSWER AS IF HEARING VOICES IS SCARY AND SEEING THINGS IS OKAY). A full list of label stimuli, target stimuli, and response options for the IRAP is provided in Table 11.

Table 11

Stimuli and Response Options of the two IRAPs employed in Experiment 6

Self IRAP		Others IRAP	
If I heard voices	If I saw things	If Other People heard voices	If Other People saw things
It would be fine I could accept it I could cope	It would be scary I could not accept it I could not cope	It would be fine They could accept it They could cope	It would be scary They could not accept it They could not cope
True	False	True	False

Procedure

The current study comprised of two IRAPs: The Self IRAP; and The Others IRAP. For all participants, there were two stages, one involving the implicit measure (the IRAP) and the second involving the explicit measures. It is important to emphasise, therefore, that each participant only completed *one IRAP*, the Self IRAP *or* the Others IRAP, hence approximately one half of the each group of participants completed each IRAP (i.e. half of the non-voice hearing control group completed the Self IRAP, while the other half completed the Others IRAP, and similarly half of the non-clinical voice hearing group completed the Self IRAP, while the other half completed the Others IRAP). Thus, the study had a between groups repeated measures design. Participants were quasi-randomly assigned to either IRAP. Prior to the IRAP, non-voice hearing and non-clinical voice hearing participants were identified using Item No. 33 of the CAPE. Non-voice hearing participants were subsequently provided with a written explanation of the phenomenon of voice hearing (see Chapter 3). And all completed the IRAP prior to the explicit measures. The verbal and automated instructions provided to participants for completing the

two IRAPs were identical to Experiment 1. For illustrative purposes, see Figure 20 for a schematic representation of the screen presentation of the IRAPs.

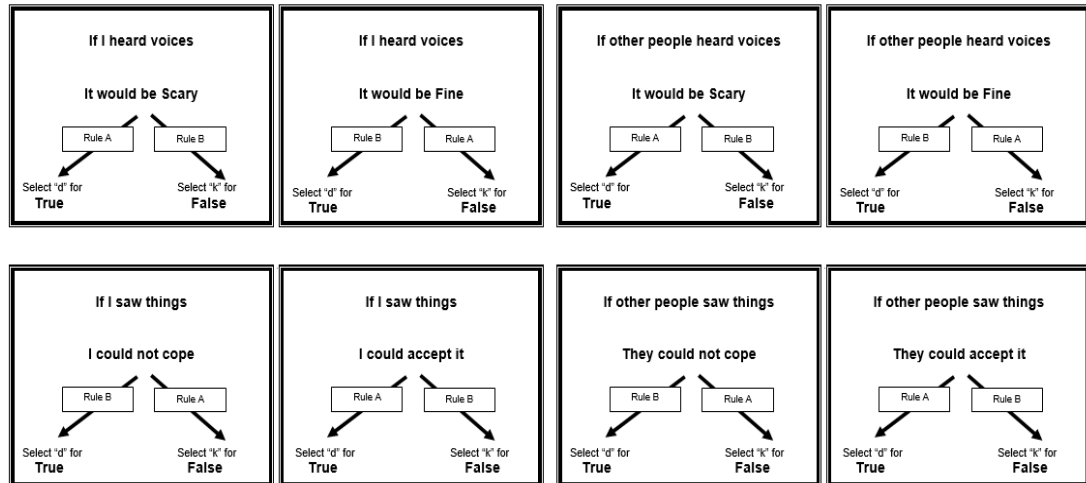


Figure 20. Schematic representation of the four trial-types presented in the Self (left) and Others (right) IRAPs in Experiment 6. The arrows and text boxes did not appear on the participant’s screen, they indicate the correct responses for Rule A and Rule B blocks of trials. **The labels used for the four trial-types in each IRAP are as follows: Voices-Scary (top-left), Voices-Okay (top-right), Visions-Scary (bottom-left) and Visions-Okay (bottom-right).**

Once participants had finished the IRAP, they completed the seven explicit measures in a pre-determined sequence (AHRS, BAVQ-R, VAAS, CAPE, AAQ-II, ATQ and SAB).

Ethical Considerations

All aspects of potential ethical considerations in Experiment 6 were identical to Experiment 5.

Results

Analytic Strategy

As in the previous chapters, all data from the visions trial-types were excluded from the analyses. To assess potential differences in the IRAP effects between the two groups in each IRAP, a two-way mixed between within ANOVA was conducted. In order to investigate the potential impact of distress of voice hearers IRAP effects, exploratory analyses of the IRAP data and its relationship with the explicit outcomes was carried out by categorising the voice hearers by level of distress level (using the AAQ-II) in distress level analyses. Correlational analyses investigated the predictive validity of the IRAP within the context of voice hearing.

Before conducting the IRAP analyses, summaries of the explicit data are provided in the section below.

Explicit Measures Data

The means and standard deviations were calculated from each participant's responses on each of the explicit measures and data are summarised in Table 12.

Table 12

Explicit Data Summary for Experiment 6

Scales	Self IRAP		Others IRAP	
	Voice Hearers <i>Mean (SD)</i>	Controls <i>Mean (SD)</i>	Voice Hearers <i>Mean (SD)</i>	Controls <i>Mean (SD)</i>
AAQ-II	20.86 (6.68)	22.00 (10.95)	27.00 (9.89)	21.70 (8.46)
CAPE				
CAPE positive dimension	3.04 (1.60)*	1.82 (0.57)*	2.47 (0.72)*	1.91 (0.47)*
CAPE depressive dimension	1.99 (1.13)	1.81 (1.10)	1.84 (1.43)*	5.68 (1.04)*
CAPE negative dimension	3.14 (1.47)	3.32 (1.15)	2.22 (0.73)	1.94 (0.62)
BAVQ-R				
Malevolence	9.57 (4.72)	-	7.09 (4.25)	-
Benevolence	13.79 (6.66)	-	12.55 (6.93)	-
Omnipotence	10.93 (4.88)	-	10.09 (4.66)	-
Emotional Resistance	8.21 (3.77)	-	6.46 (3.75)	-
Behavioural Resistance	10.50 (5.86)	-	10.55 (6.66)	-
Emotional Engagement	9.93 (6.03)	-	9.91 (6.49)	-
Behavioural Engagement	7.21 (3.40)	-	5.73 (4.41)	-
VAAS	58.07 (20.06)	-	48.00 (7.28)	-
AHRS	10.43 (8.67)	-	14.55 (6.88)	-
ATQ				
PMDC	13.14 (4.29)	12.46 (6.89)	13.18 (4.73)	13.30 (4.35)
NSNE	14.29 (6.67)	13.85 (7.03)	15.00 (5.66)	15.50 (6.45)
LSE	3.14 (1.83)	3.46 (2.15)	3.82 (2.32)	3.10 (1.37)
Helplessness	3.71 (1.90)	3.54 (2.07)	4.55 (2.42)	4.30 (1.77)
Believability	78.71 (33.03)	65.31 (32.81)	70.36 (28.03)	64.60 (22.21)
SAB	19.64 (8.75)	21.54 (7.26)	19.27 (6.97)	18.30 (5.68)

*Note. Maximum scores are: AAQ-II = 70; CAPE dimensions = 6; BAVQ-R: Mal, Ben & Omni = 18; Beh. Res. = 15, Emo. Res., Beh. Eng., Emo. Eng. = 12; VAAS = 155; AHRS = 44; ATQ: PMDC = 25; NSNE = 35; LSE = 10; Helplessness = 10; Believability = 150; SAB = 56. Missing values for explicit measures which were not administered to the control participants are denoted by "-". Significant differences denoted by * ($p < 0.05$).

For both groups, AAQ-II means revealed that participants overall had low rates of inflexibility. On the CAPE, means revealed that participants overall had low to moderate rates of psychotic-like experiences. For the ATQ, means revealed that participants overall had low to moderate rates of automatic thoughts. And on the SAB, means revealed low stigmatising attitudes toward mental illness.

For the voice hearers, BAVQ-R means revealed moderate rates of malevolence, benevolence, omnipotence, emotional resistance, emotional engagement and behavioural engagement, and high rates of behavioural resistance. On the VAAS, means revealed low overall voice acceptance. And on the AHRS, means revealed low voice severity. Notably, independent t-tests investigated potential differences across the relevant explicit measures between the voice hearing and control groups on each IRAP, and some were found (see Table 12).

IRAP Data

Scoring of the IRAP was always conducted using the standardised approach for transforming latency data into D_{IRAP} scores (see Nicholson & Barnes-Holmes, 2012a). All data from any participant which fell below 80% accuracy and above 2000ms latency on any of the six test blocks were omitted from the dataset ($N=19$). The final dataset comprised of 48 participants: 23 non-voice hearers (13 in Self IRAP and 10 in Others IRAP); and 25 non-clinical voice hearers (14 in the Self IRAP and 11 in the Others IRAP).

Between groups analyses (non-clinical voice hearers and non-voice hearers). The mean D_{IRAP} scores for the IRAP are presented in Figure 21 (visions trial-types are excluded). On the Self IRAP, on Voices-Okay, both groups showed a similar pro-voices effect. On Voices-Scary, both groups showed anti-voices effects, although the voice hearers' effect was negligible. On the Others IRAP, on Voices-Okay, both groups showed pro-voices effects, whereas on Voices-Scary, both groups showed anti-voices effects.

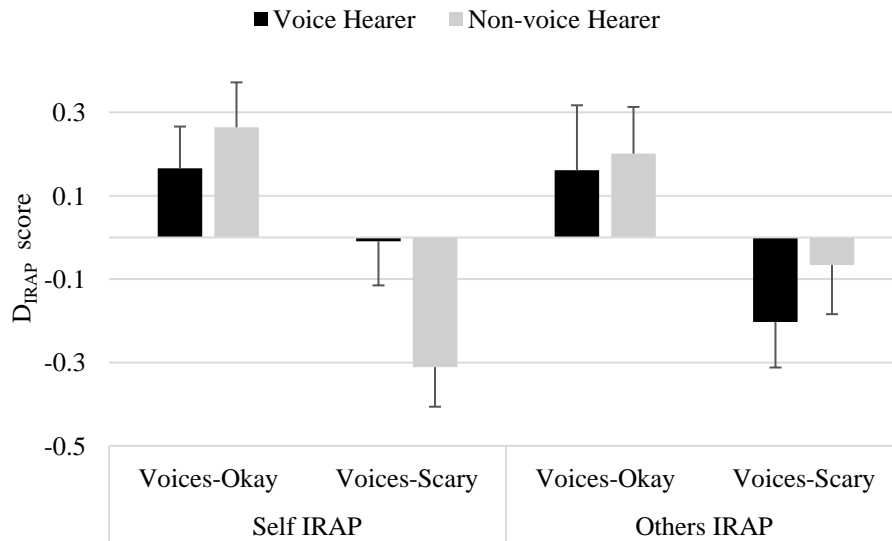


Figure 21. Mean D_{IRAP} scores for the two groups on the Self and Others IRAPs in Experiment 6. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

In order to investigate the effects of group on trial-type, a two-way mixed between within ANOVA found a main effect for trial-type ($df=1$, $F=25.884$, $p<0.0001$, $\mu^2=1.0$), and an interaction effect between trial-type, group and condition ($df=1$, $F=4.361$, $p<0.05$, $\mu^2=0.522$). Post-hoc analyses as four independent t-tests investigated potential differences between the two groups on each trial-type in each IRAP. The only differences were found on the Self IRAP on Voices-Scary ($df=25$, $t=-2.107$, $p<0.05$). Four further independent t-tests investigated potential differences between the IRAPs for each group, but found none (all $p's>0.05$). And finally, eight one-sample t-tests investigated which trial-types differed significantly from zero. For the controls, both trial-types in the Self IRAP were significant (Voices-Scary: $df=12$, $t=3.277$, $p<0.01$; Voices-Positive: $df=12$, $t=-2.442$, $p=0.05$).

Distress level analyses. Distress level analyses investigated the putative relationship between explicit distress and responding on the IRAP. For these

analyses, *the control group was excluded* and the voice hearing group was now divided along a median split on the AAQ-II.

A median split was performed to separate high ($N=11$) versus low ($N=14$) psychological inflexibility groups according to AAQ-II scores. The mean D_{IRAP} scores for each inflexibility level on the IRAP are presented in Figure 22. On the Self IRAP, on Voices-Okay, both groups showed comparable pro-voices effects. On Voices-Scary, the low inflexibility group showed anti-voices effects and the high group showed pro-voices effects. On the Others IRAP, again on Voices-Okay, both groups showed comparable pro-voices effects, however, on Voices-Scary, both groups showed anti-voices effects.

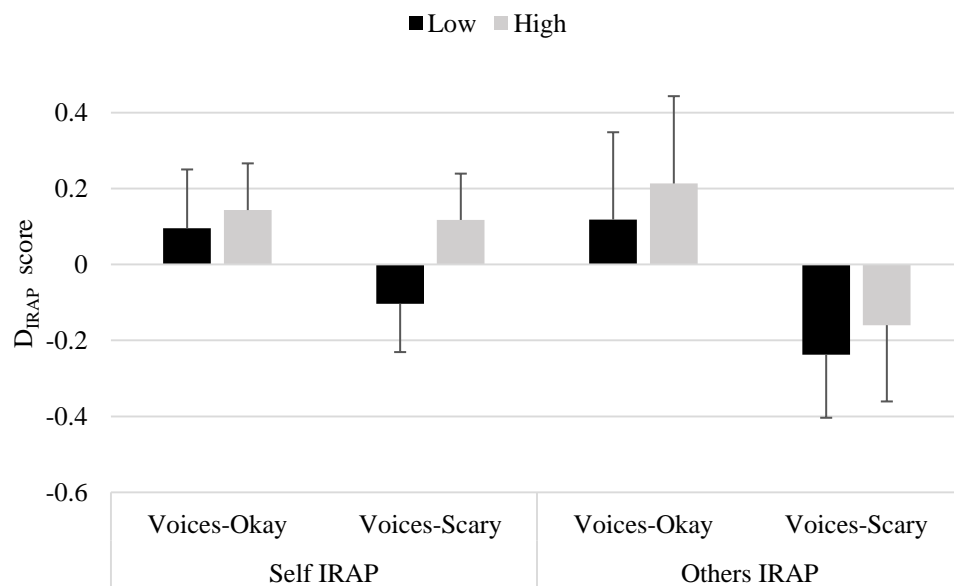


Figure 22. Mean D_{IRAP} scores for the two AAQ-II inflexibility levels on the Self and Others IRAP in Experiment 6. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

A two-way mixed between within ANOVA investigated the effects of psychological inflexibility and IRAP type on the trial-types and found a main effect for trial-type ($df=1$, $F=5.313$, $p<0.05$, $\mu^2=0.589$, all other p 's >0.05). Four

independent t-tests showed no differences between the groups on each trial-type, and eight one sample t-tests showed no significant differences for any trial-type from zero (all p 's > 0.05).

Correlations

A correlation matrix also investigated potential relationships between the IRAP effects and the explicit measures among the voice hearers in each IRAP. On the Self IRAP, a significant positive correlation was found between Voices-Okay and Depressive Frequency (CAPE: $r=0.622$, $p<0.05$). That is, the greater the depressive frequency the greater positivity toward voices.

In the Others IRAP, a significant negative correlation was found between Voices-Scary and VAAS ($r=-0.627$, $p<0.05$), that is, the more the anti-voices effect, the more acceptance of voices.

Bonferroni adjustment. Due to multiple correlational analyses conducted, a Bonferroni adjustment of the alpha level was applied ($p<0.0005$). However, using this p-value, the above correlations do not reach significance (all p 's > 0.0005). Indeed, while this adjustment is deemed necessary for statistical conservatism and to avoid the occurrence of a Type 1 error, for the purposes of the current thesis, the correlations which were significant at the 0.05 level will also be discussed.

Summary and Discussion

This experiment sought to investigate potential differences in voice hearing and control participants' evaluations of the *self* and *others* hearing voices, using the Self IRAP and the Others IRAP. Overall, hearing voices was evaluated positively by both groups on both IRAPs, although control participants were more positive on the

Self IRAP. Interestingly, control participants were also more fearful on the Self IRAP, while the voice hearers were more fearful on the Others IRAP. Indeed, both groups' positivity toward the self hearing voices was significantly different from zero, but only the control participants' *fear* was significantly different from zero. Again, the distress level analyses allowed us parse out these implicit attitudes and interestingly indicated that these positivity effects were influenced by participants' level of psychological inflexibility among the voice hearers. In summary, avoidant voice hearers were positive about their own voices, but fearful about others hearing voices. In contrast, less avoidant voice hearers were fearful about their own voices, which is consistent with the literature on avoidance where avoidance facilitates escape from fear (Luciano et al., 2013).

Again, the use of the two samples, supplemented by comparisons between the Self and Others IRAPs, allowed the precision of the IRAP to be highlighted. For example, positivity and fear on the Self IRAP were influenced by participants' own levels of psychological inflexibility. In fact, it seems highly intuitive that avoidant voice hearers would show this type of positivity while less avoidant voice hearers would show more fear, if one considers that avoidance/inflexibility facilitates escape from fear. And, this outcome is consistent with the literature on well-being and psychological avoidance. Furthermore, voice hearers who were less positive on the Self IRAP correlated with high depressive psychotic-like symptoms, which appear to complement previous evidence that there is an inverse relationship between benevolent appraisals and depression (van der Gaag et al., 2003), while those who were fearful on the Others IRAP correlated with low voice acceptance. In any case, we were now able to begin to look at some precise psychological features of voice hearers implicit attitudes to their own experiences.

Limitations

There are a number of limitations of Experiment 6 which should be reflected on to guide future research. 1. The time-point at which participants were hearing voices (i.e. past/present) was not controlled for and may have influenced the distress-level analyses. 2. The use of negatively worded target stimuli (i.e. could *not* cope, could *not* accept it) can potentially cause difficulty for participants when undertaking the IRAP, therefore, future research should try to circumvent this issue by avoiding the use of ‘not’ in target phrases. 3. The current sample comprised of a high proportion of psychology undergraduates exposed to some level of psychological training, which may account for some of the positivity, therefore, it would be interesting to replicate this study in a sample with no training in psychology.

Chapter 7
Experiment 7

Assessing fear of voices in clinical
and non-clinical voice hearers and
non-voice hearing controls

Experiment 7

Assessing Implicit Fear of Voices

Experiment 6 was the first to compare controls and voice hearers in terms of implicit fear toward voice hearing and found some interesting differences between the two groups and across the Self and Others IRAPs. In Experiment 7, we attempted to explore this implicit fear further by using a more broadly fear-based IRAP, and most notably by now including our first recruitment of clinical voice hearers. Based on previous findings, we hypothesised that at least control and non-clinical voice hearing participants would show implicit positivity and more or less fear on the IRAP. However, given that we had not previously included clinical voice hearers, it was difficult to predict what effects would be recorded with this group in terms of positivity or fear, however we hypothesised less positivity in this regard. It was equally hard to hypothesise about whether this group would show differential effects to the other two. In short, it was of particular interest to us to see whether, at various levels, non-clinical voice hearers would be more similar in their responses to controls or to clinical voice hearers. This was a key aim of Experiment 7.

Method

Setting

The current study was conducted in two locations. The non-clinical voice hearers and the non-voice hearers participated in an experimental cubicle at the Department of Psychology, MU. The clinical voice hearers participated in a research room at a psychiatric facility in Ireland. All participation was on an individual basis. For the non-clinical voice hearers and the non-voice hearers, the experimenter

interacted with participants only during instructional phases of the IRAP, and remained seated behind participants at all other times. On average, experimental sessions lasted between 30 and 60 minutes and all participation was completed in one session. For the clinical voice hearers, it was necessary for the experimenter to interact with participants during all phases of the study. On average, experimental sessions lasted between 1.5 and 4 hours (with regular breaks as requested) and all participation was completed in two to eight sessions.

Participants

The current study involved three groups of participants. One group were categorised as clinical voice hearers, another categorised as non-clinical voice hearers and the third group comprised of a non-voice hearing control group. Ten clinical voice hearers (independently diagnosed but presenting with various diagnoses: Schizophrenia $N=6$, Bipolar Disorder $N=2$, and Depression with Psychotic Features $N=2$) were recruited from St. Patrick's Mental Health Services Hospital in Dublin, Ireland, six of these were male and four were female. Seventeen non-clinical voice hearers were identified as such using current screening methods from a general sample of undergraduate students recruited from MU. Nine of these were male and eight were female. Twenty-five non-voice hearing individuals were identified from the same general sample of undergraduate students recruited from MU. Twelve of these were male and 13 were female. In total, the study involved 37 males and 25 females, with an age range of 18 to 37 years, with a mean age of 22.18 years and a standard deviation of 4.09 years.

Materials

Explicit measures. Three broad categories of explicit measures were administered. The first series of measures assessed voice hearing (AHRS, BAVQ-R and VAAS) and psychosis (CAPE). The second set of measures more broadly assessed general psychological well-being (AAQ-II and ATQ). The third measured stigma toward mental health difficulties (SAB). See Chapters 2, 3 and 5 for outlines of all explicit measures.

The IRAP. The IRAP was administered on a standard desktop computer. This was used to present the instructions and stimuli, and to record responses. The current study involved the Fear IRAP that assessed implicit bias regarding the fear of voices (see Chapter 4 for outline of the Fear IRAP).

Procedure

The current study comprised of two stages, one involving the implicit measure (the IRAP), and the second involving the explicit measures. Thus, the study had a between groups repeated measures design. Prior to the IRAP, non-voice hearing and non-clinical voice hearing participants were identified using Item No. 33 of the CAPE. Non-voice hearing participants were subsequently provided with a written explanation of the phenomenon of voice hearing (see Chapter 3).

All participants completed the IRAP prior to the explicit measures. *The IRAP employed here was the Fear IRAP and was identical to that presented in Chapter 4.* The verbal and automated instructions provided to participants for completing the IRAP were identical to Experiment 1. Once participants had finished the IRAP, they completed the seven explicit measures in a pre-determined sequence (AHRS, BAVQ-R, VAAS, CAPE, AAQ-II, ATQ and SAB).

Ethical Considerations

This experiment included the addition of a clinical voice hearing sample and the steps taken to circumvent any potential ethical concerns can be summarised as follows. 1. This study received prior ethical approval from Maynooth University Ethics Committee (Approval date 18/06/13) and St. Patrick's University Hospital (Ref. No. 25/13). 2. All participants who are not drawn from a clinical sample received the information sheet and this is followed immediately by the consent forms (with participation potentially immediate thereafter). In contrast, all participants drawn from a clinical sample receive the information sheet followed by a delay of one week minimum prior to receiving the consent forms (with participation potentially immediate thereafter). 3. All clinical participants were offered the opportunity to take a 10-minute break (approximately) between each stage of their experiment. 4. All participation from the clinical sample was referred by a senior clinical psychologist. 5. All clinical participants were given details of a Clinical Psychologist who could provide post-research support for potential issues that emerged from participation, however none did. All other aspects of potential ethical considerations in Experiment 7 were identical to Experiment 5.

Results

Analytic Strategy

As in the previous chapters, all data from the visions trial-types were excluded from the analyses. To assess potential differences in IRAP effects among groups of participants, a mixed between within ANOVA was conducted. The literature has suggested that categorising voice hearing using a clinical versus non-clinical distinction does not always yield meaningful results (Longden, Corstens,

Escher, & Romme, 2012). Therefore, a more appropriate means of analysing the IRAP data and its relationship with the explicit outcomes was to categorise the voice hearers by level of psychotic-like symptoms (using the CAPE), by distress level (using the AAQ-II), and by perceived controllability of voices analyses (using the AHRs). Correlational analyses were then used to investigate the predictive validity of the IRAP within the context of voice hearing. It was hypothesised that voice hearers with greater levels of psychotic-like symptoms and distress and those with less perceived control over voices, would show less implicit positivity towards voices.

Before conducting the IRAP analyses, summaries of the explicit data are provided in the section below.

Explicit Measures Data

The means and standard deviations were calculated from each participant's responses on each of the explicit measures and data are summarised in Table 13.

Table 13

Explicit Data Summary for Experiment 7

Scales	Clinical Voice Hearers	Non-Clinical Voice Hearers	Controls	ANOVA/ T-Test
	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>F/T</i>
AAQ-II	29.00 (8.66)	22.77 (7.48)	19.65 (8.37)	1.964
CAPE				
CAPE positive dimension	1.95 (1.77)	2.23 (0.84)	2.05 (0.58)	.273
CAPE depressive dimension	1.75 (1.89)	3.13 (1.29)	3.27 (1.21)	1.774
CAPE negative dimension	1.43 (1.25)	3.26 (1.56)	2.57 (0.76)	3.348*
BAVQ-R				
Malevolence	17.33 (10.21)	6.71 (5.03)	-	-3.725**
Benevolence	12.00 (10.58)	11.12 (8.02)	-	-.169
Omnipotence	15.67 (3.51)	8.00 (5.71)	-	-2.222*
Emotional Resistance	11.33 (4.51)	6.24 (5.54)	-	-1.498
Behavioural Resistance	14.67 (5.13)	10.00 (6.71)	-	-1.137
Emotional Engagement	7.33 (4.51)	7.41 (5.10)	-	.025
Behavioural Engagement	5.67 (3.51)	5.88 (4.21)	-	.083
VAAS	61.67 (51.08)	47.65 (15.72)	-	-.992
AHRS	39.33 (10.97)	11.12 (8.05)	-	-5.351***
ATQ				
PMDC	7.00 (7.00)	11.06 (5.74)	10.71 (5.31)	.673
NSNE	8.67 (8.08)	11.59 (7.04)	12.00 (6.02)	.321
LSE	2.00 (2.00)	3.24 (2.73)	2.29 (1.36)	.992
Helplessness	2.33 (2.52)	3.35 (2.26)	3.18 (1.78)	.311
Believability	41.00 (40.51)	70.35 (42.14)	73.76 (43.37)	.759
DASS				
Depression	41.67 (29.40)	28.94 (17.79)	29.18 (21.60)	.523
Anxiety	15.33 (14.19)	6.82 (6.25)	7.88 (8.90)	1.371
Stress	12.00 (9.17)	7.29 (7.61)	8.24 (8.09)	.452
	14.33 (8.62)	15.18 (8.22)	13.41 (9.82)	.162
SAB	8.33 (14.43)	19.65 (5.57)	21.47 (5.67)	5.246**

*Note. Maximum scores are: AAQ-II = 70; CAPE dimensions = 6; BAVQ-R: Mal, Ben & Omni = 18; Beh. Res. = 15, Emo. Res., Beh. Eng., Emo. Eng. = 12; VAAS = 155; AHRS = 44; ATQ: PMDC = 25; NSNE = 35; LSE = 10; Helplessness = 10; Believability = 150; DASS total = 126; DASS subscales = 42; SAB = 56. Missing values for explicit measures which were not administered to the control participants are denoted by "-". Significant effects denoted by * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$).

For all groups, AAQ-II means revealed that participants overall had low rates of inflexibility. On the CAPE, means revealed that participants overall had low to moderate rates of psychotic-like experiences. For the ATQ, means revealed that participants overall had low to moderate rates of automatic thoughts, the lowest rates

were found in the clinical voice hearing group. And on the SAB, means revealed low stigmatising attitudes toward mental illness.

For the non-clinical voice hearers, BAVQ-R means revealed moderate rates of malevolence, benevolence, omnipotence, emotional resistance, emotional engagement and behavioural engagement, and high rates behavioural resistance. For the clinical voice hearers, higher rates were found in all subscales, except for behavioural and emotional engagement with voices. On the VAAS, means revealed higher rates of overall voice acceptance in the clinical voice hearers, when compared to their non-clinical counterparts. And on the AHRS, means revealed greater voice severity in the clinical voice hearers, when compared to the non-clinical group. One-way ANOVAs were conducted to investigate potential differences between the three groups of participants, and independent t-tests were conducted to investigate potential differences between the two groups of voice hearers, and some were found (see Table 13).

IRAP Data

Scoring of the IRAP was always conducted using the standardised approach for transforming latency data into D_{IRAP} scores (see Nicholson & Barnes-Holmes, 2012a). All data from any participant which fell below 80% accuracy and above 2000ms latency on any of the six test blocks were omitted from the dataset ($N=15$). The final dataset comprised of 37 participants: 17 non-voice hearers (10 male and 7 female); 17 non-clinical voice hearers (9 male and 8 female); and 3 clinical voice hearers (1 male and 2 female).

Between groups analyses (clinical, non-clinical and controls). The mean D_{IRAP} scores for the three groups on the IRAP are presented in Figure 23 (visions

trial-types are excluded). On Voices-Okay, the controls and non-clinical voice hearers showed pro-voices effects, with the clinical voice hearers showed anti-voices effects. On the Voices-Scary trial-type, all three groups showed anti-voices effects.

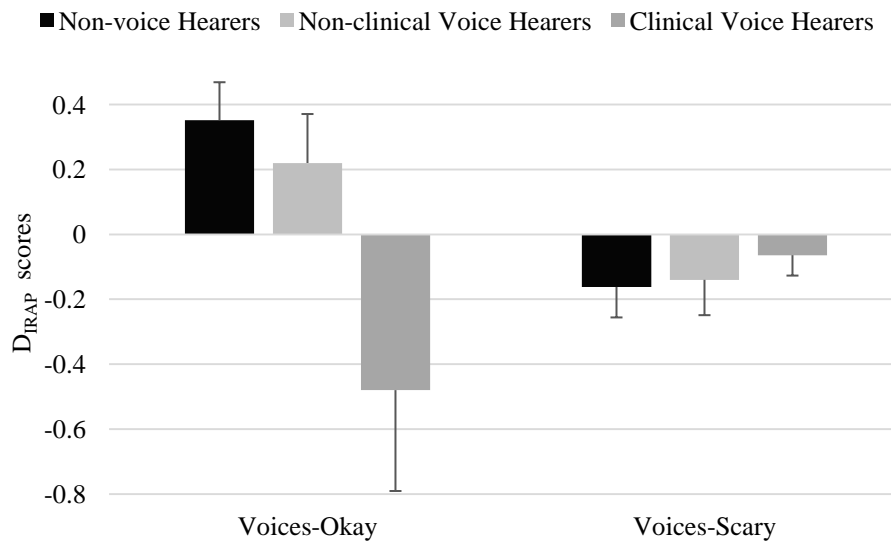


Figure 23. Mean D_{IRAP} scores for the three groups on the Fear IRAP in Experiment 7. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

To investigate the effects of group on trial-type, a mixed between within ANOVA found no significant main effects (all p 's > 0.05), but did find an interaction effect between group and trial-type ($F=3.660$, $p<0.05$, $\eta^2=0.631$). Post-hoc analyses, as six independent t-tests, determined potential differences among the groups on individual trial-types, and found one difference between the clinical group and the non-voice hearing group on Voices-Okay ($df=18$, $t=-2.719$, $p<0.05$, all other p 's > 0.05). Six one sample t-tests investigated whether the D_{IRAP} trial-type effects differed significantly from zero. The only significant difference emerged was among the non-voice hearers on Voices-Okay ($df=16$, $t=-3.010$, $p<0.01$, all other p 's > 0.05).

CAPE analyses. Data from the clinical and non-clinical voice hearers was split into two groups using a median split on the positive dimension subscale, thus creating high vs. low CAPE groups. The median split was conducted as only three participants fell over the recommended cut-off for being at-risk of psychosis (Mossaheb et al., 2012) . The mean D_{IRAP} scores for the high ($N=11$) and low ($N=19$) positive dimension CAPE scores on the IRAP are presented in Figure 24. For Voices-Okay, the low CAPE group showed pro-voices effects, while the high group showed a negligible effect. For Voices-Scary, both showed anti-voices effects (the larger of which was in the high group).

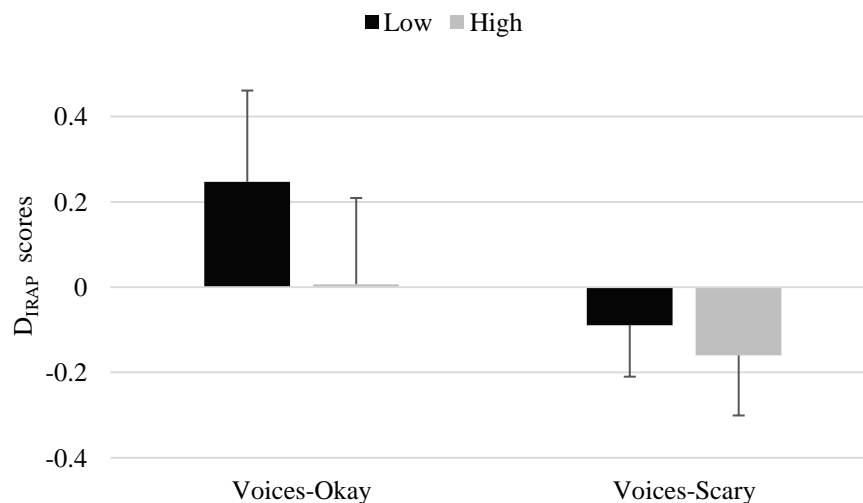


Figure 24. Mean D_{IRAP} scores on the Fear IRAP for voice hearers and high and low CAPE groups in Experiment 7. **Again, positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

A mixed between within ANOVA produced no main or interaction effects (all $p's > 0.05$). Exploratory trial-type analyses, as two independent t-tests, investigated potential differences across groups on each trial-type, but no significant differences were found (all $p's > 0.05$). Again, four one-sample t-tests investigated

whether the effects in each group differed significantly from zero, but none did (all $p's > 0.05$).

Distress level analyses. Distress level analyses investigated the putative relationship between explicit distress and responding on the IRAP. For these analyses, the control group was excluded, the two voice hearing groups were combined, and this single group was now divided along a median split on the AAQ-II.

AAQ-II. A median split was performed to separate high ($N=10$) versus low ($N=10$) psychological inflexibility groups according to AAQ-II scores. The mean D_{IRAP} scores for each inflexibility level on the IRAP are presented in Figure 25. On Voices-Okay, both groups showed comparable pro-voices effects and on Voices-Scary, both groups show anti-voices effects (the larger effect was in the high group). A mixed between within ANOVA investigated the effects of psychological inflexibility on the trial-types, but no effects were found (all $p's > 0.05$). Two independent t-tests showed no differences between the groups, and four one sample t-tests showed no significant differences for any trial-type relative to zero (all $p's > 0.05$).

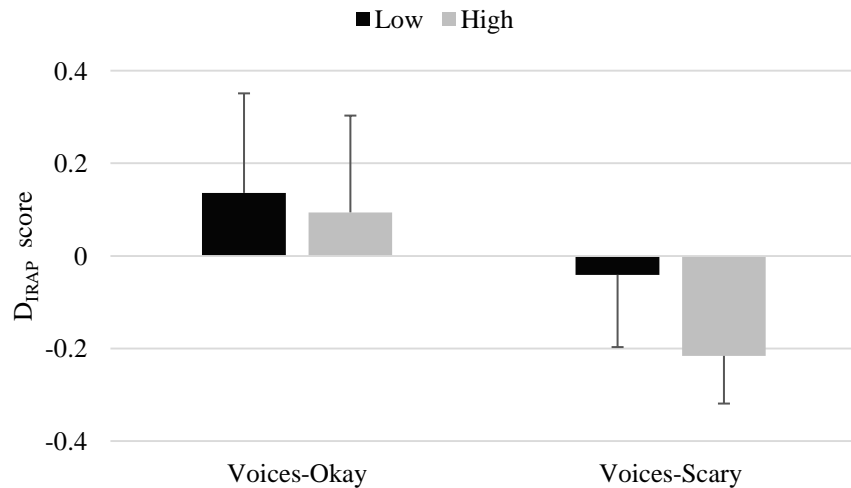


Figure 25. Mean D_{IRAP} scores for the two AAQ-II inflexibility levels on the Fear IRAP in Experiment 7. **Again, positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

Perceived controllability analyses. The mean D_{IRAP} scores for the four perceived controllability dimensions as measured by the AHRS (i.e. full control [$N=9$], some-majority of time [$N=5$], some-occasionally [$N=4$], and no control [$N=2$]) on the IRAP are presented in Figure 26. On Voices-Okay, the some-majority and no control groups also showed anti-voices effects, whereas the full control and some-occasionally groups showed pro-voices effects. On Voices-Scary, three of the groups (excluding some-occasionally) showed anti-voices effects, with the largest recorded in the some-majority group.

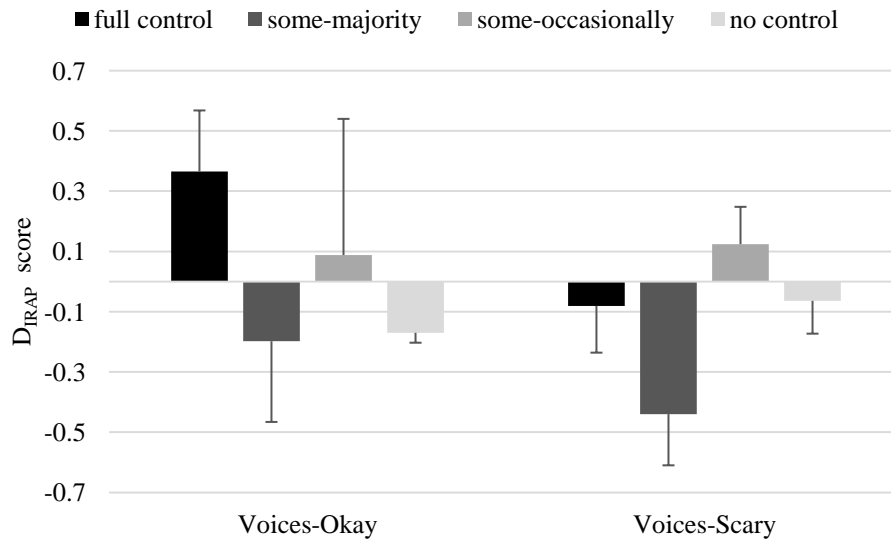


Figure 26. Mean D_{IRAP} scores for AHRS perceived controllability on the Fear IRAP in Experiment 7. **Again, positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

A mixed between within ANOVA investigated potential differences among the perceived controllability groups and IRAP effects, but no effects were found (all $p's > 0.05$). Twelve independent t-tests investigated group differences on each trial-type and only some-majority and some-occasional control differed significantly on Voices-Scary ($df=7, t=2.553, p < 0.05$). Eight one-sample t-tests investigated whether the D_{IRAP} effects differed significantly from zero, but none did (all $p's > 0.05$).

Correlations

A correlation matrix also investigated potential relationships between the IRAP effects and the explicit measures among the voice hearers. Significant positive correlations were only found between: voice severity (on the AHRS) and Voices-Scary ($r=0.435, p < 0.05$), that is, the greater the anti-voices effect, the greater voice severity; and also Voices-Okay and behavioural resistance (BAVQ-R; $r=0.440$,

$p < 0.05$), that is, the greater the pro-voices effect, the greater the behavioural resistance of voices.

Bonferroni adjustment. Due to multiple correlational analyses conducted, a Bonferroni adjustment of the alpha level was applied ($p < 0.0005$). However, using this p-value, the above correlations do not reach significance (all p 's > 0.0005). Indeed, while this adjustment is deemed necessary for statistical conservatism and to avoid the occurrence of a Type 1 error, for the purposes of the current thesis, the correlations which were significant at the 0.05 level will also be discussed.

Summary and Discussion

This was the first study to involve clinical voice hearers and compared these with controls and non-clinical voice hearers on the Fear IRAP. As expected, controls and non-clinical voice hearers showed implicit positivity and only marginal fear on the IRAP. But, most interestingly, the clinical voice hearers were much less positive in this regard. And only the control participants' positivity was significantly different from zero in this instance. The distress level analyses showed that voice hearers who were low on psychotic-like experiences *and* inflexibility were positive, while those who were high on these experiences were less positive and more fearful. And voice hearers who reported *full* control evaluated voices most positively, while those with *some or no* control were most fearful. The correlations also showed that less implicit fear correlated with higher voice severity on the AHRS, which appears to contradict the finding that appraisals are independent of voice severity (e.g. van der Gaag et al., 2003). Once again, these effects supported those we had observed previously in the thesis, but critically showed that the clinical voice hearers were implicitly less positive than the other two groups.

Limitations

There are a number of limitations of Experiment 7 which should be reflected on to guide future research. 1. The time-point at which participants were hearing voices (i.e. past/present) was not controlled for and may have influenced the distress-level analyses. 2. The low N of clinical voice hearers. Future studies should try to include a larger sample size. 3. The use of inferential statistics in low N analyses (i.e. the distress level analyses) does not allow researchers to observe the differences or effects which may exist. Future research should try to include larger N 's for the distress level analyses in order to better examine this nuanced relationship between one's own level of distress and implicit attitudes.

Chapter 8

Experiment 8

Assessing the valence and acceptance of voices in clinical and non-clinical voice hearers and non-voice hearing controls

Experiment 8

Assessing Implicit Valence and Acceptance of Voices

Some authors have suggested that non-clinical and clinical voice hearers may be distinguished in terms of the extent to which they perceive control over their voices, accept or avoid voices, and the amount of negatively valenced content of voices (Brett et al., 2007; Daalman et al., 2011). Hence, Experiment 8 sought to explore the potentially different reactions of these two groups to voices, and in doing so we attempted to parse out emotional versus behavioural responses. In a Valence IRAP, we contrasted positive and negative reactions to voices, and an Acceptance IRAP contrasted avoidance and acceptance of positive and negative voices.

Based on previous findings, we hypothesised that at least non-voice hearing controls and non-clinical voice hearing participants would show implicit positivity, less of which would likely be observed in the clinical group. Because we had not previously targeted behavioural responses (i.e. acceptance) on the IRAP, it was difficult to hypothesise how the three groups would respond in this regard. However, based our previous findings on explicit inflexibility, and differences we had observed between non-clinical and clinical voice hearers, we hypothesised that clinical voice hearers may have less acceptance of voices than non-clinical voice hearers. This was the key aim of Experiment 8.

Method

Setting

The current study was conducted in two locations. The non-clinical voice hearers and the non-voice hearing controls participated in an experimental cubicle at

the Department of Psychology, MU. The clinical voice hearers participated in a research room in a psychiatric facility in The Netherlands. All participation was on an individual basis. For the non-clinical voice hearers and the non-voice hearers, the experimenter interacted with participants only during instructional phases of the IRAP, and remained seated behind participants at all other times. On average, experimental sessions lasted between 30 and 60 minutes and all participation was completed in one session. For the clinical voice hearers, it was necessary for the experimenter to interact with participants during all phases of the experiment. On average, experimental sessions lasted between 1.5 and 4 hours (with regular breaks as requested) and all of participation was completed in two to eight sessions.

Participants

The current study involved three groups of participants. One group were categorised as clinical voice hearers, another categorised as non-clinical voice hearers and the third group comprised of a non-voice hearing control group. Fourteen clinical voice hearers (independently diagnosed but presenting with various diagnoses: Schizophrenia $N=8$, Schizophreniform Disorder $N=1$, Bipolar Disorder $N=3$, and Depression with Psychotic Features $N=2$) were recruited from the Vincent van Gogh Psychiatric facility in Venray, The Netherlands, six of these were male and eight were female. Thirty-two non-clinical voice hearers were identified as such using current screening methods from a general sample of undergraduate students recruited from MU. Sixteen of these were male and 16 were female. Twenty-seven non-voice hearing individuals were identified from the same general sample of undergraduate students recruited from MU. Twelve of these were male and 15 were female. In total, the study involved 55 participants: 27 males and 28 females, with an

age range of 18 to 38 years and a mean age of 25.18 years and a standard deviation of 2.19 years.

Materials

Explicit measures. Two broad categories of explicit measures were administered. The first series of measures assessed voice hearing (AHRS, BAVQ-R and VASS) and psychosis (CAPE). The second set of measures more broadly assessed general psychological well-being (AAQ-II, ATQ and CAPE). See Chapters 2, 3 and 5 for outlines of all explicit measures.

The IRAP. The IRAP was administered on a standard desktop computer. This was used to present the instructions and stimuli, and to record responses. The current study involved two IRAPs, one that assessed implicit appraisals regarding the valence of voices (referred to as the Valence IRAP) and another that assessed the acceptance of voices (referred to as the Acceptance IRAP). For Dutch participants, the IRAP stimuli were presented in Dutch (translated by a research team at the psychiatric institution in Venray), and for Irish participants, the IRAP stimuli were presented in English (see Table 14). Stimuli were selected by a group of twelve researchers and clinicians who had direct experience of working with voice hearers.

The Valence IRAP contrasted hearing voices with seeing things, using the labels THE VOICES I HEAR and THE THINGS I SEE. Each trial-type presented one of these two types of category labels. These were accompanied by one of eight positive (e.g. ARE HELPFUL) or eight negative target stimuli (e.g. ARE SCARY). The screen also presented two response options, TRUE and FALSE. Before each block of trials, the screen presented one of two rules for responding (i.e. PLEASE ANSWER AS IF HEARING VOICES IS NEGATIVE AND SEEING THINGS IS

POSITIVE or PLEASE ANSWER AS IF HEARING VOICES IS POSITIVE AND SEEING THINGS IS NEGATIVE).

The Acceptance IRAP contrasted positive and negative voices, using the labels IF MY VOICES ARE PLEASANT and IF MY VOICES ARE ANNOYING. Each trial-type presented one of these two types of category labels. These were accompanied by one of six avoidance-based (e.g. I BLOCK THEM OUT) or six acceptance-based target stimuli (e.g. I CHERISH THEM). Again, the screen also presented the response options, TRUE and FALSE. Before each block of trials, the screen presented one of two rules for responding (i.e. PLEASE ANSWER AS IF PLEASANT VOICES ARE POSITIVE AND ANNOYING VOICES ARE NEGATIVE or PLEASE ANSWER AS IF PLEASANT VOICES ARE NEGATIVE AND ANNOYING VOICES ARE POSITIVE). A full list of label stimuli, target stimuli, and response options for each IRAP in both English and Dutch is provided in Table 14.

Table 14

Stimuli and Response Options in the English and Dutch of the two IRAPs employed in Experiment 8

Valence IRAP ³				Acceptance IRAP			
The voices I hear	De stemmen die ik hoor	The things I see	De dingen die ik zie	If my voices are pleasant I	Als mijn stemmen prettig zijn dan	If my voices are annoying I	Als mijn stemmen prettig zijn dan
Lift me	Geven mij zelfvertrouwen	Devalue me	Zijn tegen mij	Welcome them	Verwelkom ik ze	Block them out	Druk ik ze weg
Give me guidance	Zijn vriendelijk	Are frightening	Zijn irritant	Try to keep them	Koester ik ze	Ignore them	Zoek ik afleiding
Help me	Zijn behulpzaam	Are scary	Halen mij naar beneden	Accept them	Accepteer ik ze	Suppress them	Leg ik ze het zwijgen op
Are kind	Geven mij richting	Are against me	Onderdrukken mij	Listen to them	Luister ik naar ze	Try to stop them	Negeer ik ze
Calm me	Zijn kalmerend	Annoy me	Zijn gevaarlijk	Cherish them	Wil ik ze houden	Abstract myself	Probeer ik ze te stoppen
Are funny	Zijn grappig	Are a hindrance	Maken mij van streek	Am open to them	Sta ik hier voor open	Shut them up	Sluit ik mij hiervoor af
Are my friend	Zijn mijn vrienden	Frustrate me	Frustreren mij				
Feel good	Geven mij een goed gevoel	Undermine me	Zijn eng				
True/Waar				False/Niet Waar			

³ Note that eight target stimuli remains within the recommended range of target stimuli for the IRAP.

Procedure

The current study comprised of two stages, one involving the implicit measure (the two IRAPs), and the second involving the explicit measures. Thus, the study had a between groups repeated measures design. Prior to the two voice hearing IRAPs (i.e. the Valence IRAP and the Acceptance IRAP), the non-voice hearers and the non-clinical voice hearers were identified using Item No. 33 of the CAPE. Non-voice hearing participants were subsequently provided with a written explanation of the phenomenon of voice hearing (see Chapter 3), and for the purposes of the current study were asked to imagine that they had these experiences. Clinical voice hearers were identified as such through their consultant psychiatrist/clinical psychologist as having a history of, or were currently hearing voices.

All participants completed the Valence IRAP first, followed by the Acceptance IRAP. All participants completed the IRAPs prior to the explicit measures. The verbal and automated instructions provided to participants for completing the two IRAPs were identical to Experiment 1. The only presentation feature that distinguished the current IRAP from those presented in Experiment 1 is that there were 32 trials in each block on the Valence IRAP. This change resulted from the presentation of eight target stimuli rather than six. For illustrative purposes, see Figure 27 for a schematic representation of the screen presentation of the IRAPs.

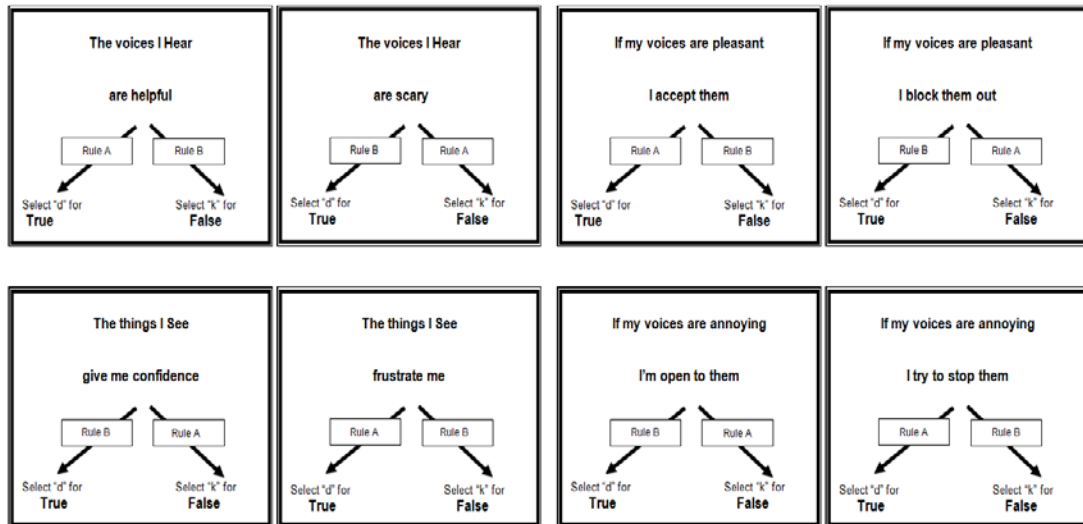


Figure 27. Schematic representation of the four trial-types presented in the Valence (left) and the Acceptance (right) IRAPs presented in Experiment 8. The arrows and text boxes did not appear on the participant’s screen, they indicate the correct responses for Rule A and Rule B blocks of trials. **The labels used for the four trial-types in the Valence IRAP (left) are as follows: voices-positive (top-left), voices-negative (top-right), visions-positive (bottom-left) and visions-negative (bottom-right). And the labels used for the four trial-types in the Acceptance IRAP (right) are as follows: pleasant-accept (top-left), pleasant-avoid (top-right), annoying-accept (bottom-left) and annoying-avoid (bottom-right).**

Once participants finished both IRAPs, they completed the seven explicit measures in a pre-determined sequence.

Ethical Considerations

This study received prior ethical approval from Maastricht University Ethics Committee (Approval date: 18/06/13) and the Vincent van Gogh Institute for Psychiatry in Venray (Ref No. UT/ec/UI3000077), The Netherlands. All other aspects of potential ethical considerations in Experiment 8 and the procedures conducted to circumvent these were identical to Experiment 7.

Results

Analytic Strategy

As in the previous chapters, all data from the visions trial-types were excluded from the analyses. To assess potential differences in the IRAP effects among groups of participants, two mixed between within ANOVAs were conducted for each IRAP. As before, the literature has suggested that categorising voice hearing using a clinical versus non-clinical distinction does not always yield meaningful results (Longden et al., 2012). Therefore, a more appropriate means of analysing the IRAP data and its relationship with the explicit outcomes was to categorise the voice hearers by distress level (using the DASS and the AAQ-II) and by perceived controllability of voices analyses (using the AHRS). Correlational analyses investigated the predictive validity of the IRAP within the context of voice hearing.

Before conducting the IRAP analyses, summaries of the explicit data are summarised in the section below.

Explicit Measures Data

The means and standard deviations were calculated from each participant's responses on each of the explicit measures and data are summarised in Table 15.

Table 15

Explicit Data Summary for Experiment 8

Scales	Clinical Voice Hearers	Non-Clinical Voice Hearers	Controls	ANOVA/T-Test
	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>F</i>
AAQ-II	46.67 (9.10)	27.28 (15.01)	23.95 (10.59)	8.292***
CAPE				
CAPE positive dimension	13.37 (10.433)	11.39 (11.33)	20.32 (6.18)	19.099***
CAPE depressive dimension	7.84 (5.05)	6.83 (7.07)	15.71 (5.7)	25.923***
CAPE negative dimension	10.77 (8.12)	9.69 (10.44)	23.59 (8.67)	26.875***
BAVQ-R				
Malevolence	4.67 (4.12)	6.04 (3.98)	-	-.875
Benevolence	8.56 (5.73)	8.96 (5.50)	-	-.185
Omnipotence	6.56 (3.36)	7.33 (3.45)	-	-.581
Resistance	10.67 (6.46)	13.83 (6.75)	-	-.165
Engagement	9.67 (6.75)	10.08 (6.34)	-	-1.213
VAAS	42.00 (9.54)	56.13 (14.71)	-	-1.087
AHRS	27.78 (12.07)	19.25 (11.05)	-	1.927
ATQ				
Frequency	76.00 (29.46)	46.04 (24.80)	34.27 (14.31)	12.023***
Believability	76.67 (28.24)	47.63 (28.57)	35.95 (14.43)	9.322**
DASS				
Depression	18.78 (12.01)	29.38 (17.36)	37.68 (12.93)	5.359**
Anxiety	6.78 (6.53)	9.38 (6.53)	12.14 (4.82)	2.917
Stress	4.89 (3.10)	9.04 (6.01)	11.64 (4.63)	5.668**
	7.11 (4.28)	10.96 (5.58)	13.91 (4.84)	5.909**

*Note. Maximum scores are: AAQ-II = 70; CAPE dimensions = 6; BAVQ-R: Mal, Ben & Omni = 18; Resistance = 27, Engagement = 24; VAAS = 155; AHRS = 44; ATQ: Frequency = 150; Believability = 150; DASS total = 126; DASS subscales = 42. Missing values for explicit measures which were not administered to the control participants are denoted by “-”. Significant effected denoted by * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$).

Overall, for the AAQ-II, the non-clinical and control groups had low levels of inflexibility, whereas the clinical voice hearers had moderate levels of inflexibility. Surprisingly, the highest score in each dimension of the CAPE were recorded in the control group. For the voice hearing measures, both voice hearing groups had low to moderate levels of voice hearing dimensions on the BAVQ-R. On the VAAS, the non-clinical voice hearers had slightly higher levels of voice acceptance than the

clinical voice hearers. And on the AHRS, the clinical voice hearers had greater voice severity than the non-clinical group. On the ATQ, the clinical voice hearers had higher levels of automatic thoughts than the other two groups. And finally, all groups showed low levels of distress on the DASS, with the highest found in the control group. One-way ANOVAs were conducted to investigate potential differences between the three groups of participants, and independent t-tests were conducted to investigate potential differences between the two groups of voice hearers, and some were found (see Table 15).

Valence IRAP Data

Scoring of the IRAP was always conducted using the standardised approach for transforming latency data into D_{IRAP} scores (see Nicholson & Barnes-Holmes, 2012a). All data from any participant which fell below 80% accuracy and above 2000ms latency on any of the six test blocks were omitted from the dataset (N=18). The final dataset comprised of 55 participants: 27 males and 28 females (Clinical voice hearers= 9, four male and five female, Non-clinical voice hearers= 24, 13 male and 11 female, and Non-voice hearers= 22, 10 male and 12 were female).

Between groups analyses (clinical, non-clinical and controls). The mean D_{IRAP} scores for the three groups on the Valence IRAP are presented in Figure 28 (visions trial-types are excluded). On Voices-Positive, clinical and non-clinical voice hearers showed anti-voices effects, but the controls showed pro-voices effects. And a similar pattern emerged on Voices-Negative, where only the clinical and non-clinical voice hearers showed anti-voices effects.

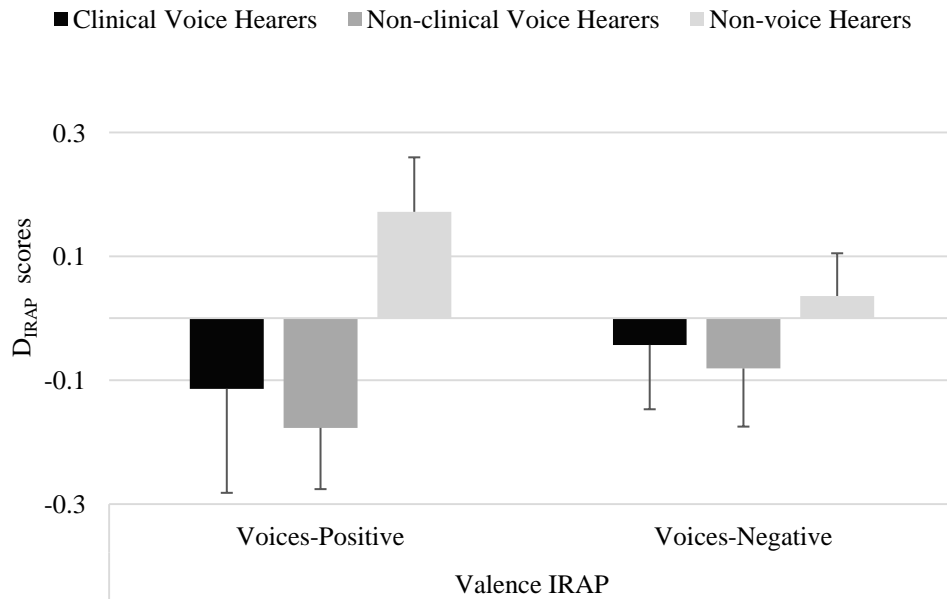


Figure 28. Mean D_{IRAP} scores for the three groups on the Valence IRAP in Experiment 8. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

To investigate the effects of group on trial-type, a mixed between within ANOVA found no main or interaction effects (all p 's > 0.05). Exploratory trial-type analyses, in the form of six independent t-tests, determined potential differences among the groups on individual trial-types, and recorded only one difference between the non-clinical group and the non-voice hearing group on Voices-Positive ($df=44$, $t=-2.699$, $p<0.01$, all other p 's > 0.05). Six one sample t-tests investigated whether the D_{IRAP} trial-type effects differed significantly from zero, but none did (all p 's > .05).

Distress level analyses. Distress level analyses investigated the putative relationship between explicit level of distress and responding on the Valence IRAP. For these analyses, the control group was excluded, the two voice hearing groups were combined, and this single group was now divided along the clinical categories of the DASS and a median split on the AAQ-II.

DASS. The overall DASS scores were separated along the clinical guidelines specified by the measure: mild ($N=11$), moderate ($N=10$), severe and extreme ($N=12$), the latter two were combined in the current analyses. The mean D_{IRAP} scores for the three distress levels on the Valence IRAP are presented in Figure 29. On voice-positive, all groups showed anti-voices effects. Similarly, on Voices-Negative, the mild and moderate groups showed anti-voices effects, while the severely group showed marginal pro-voices effects. A mixed between within ANOVA investigated the effects of distress on IRAP trial-types, but found no significant effects (all p 's >0.05). Exploratory trial-type analyses, in the form of six independent t-tests found no significant differences (all p 's >0.05). However, six one-sample t-tests showed that for the moderately distressed groups, the D_{IRAP} effect was significant for Voices-Positive ($df=9$, $t=-2.349$, $p<0.05$). All other effects were non-significant (all p 's >0.05).

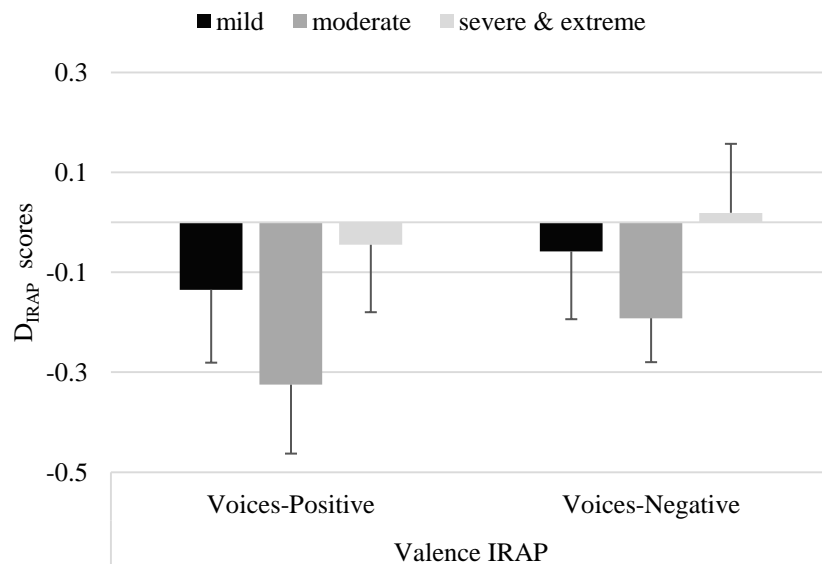


Figure 29. Mean D_{IRAP} scores for the three DASS distress groups on the Valence IRAP in Experiment 8. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

AAQ-II. A median split was performed to separate high ($N=16$) versus low ($N=17$) psychological inflexibility groups according to AAQ-II scores (there are no clinical cut-offs recommended for this measure). The mean D_{IRAP} scores for each inflexibility level on the Valence IRAP are presented in Figure 30. On Voices-Positive and Voices-Negative, both groups showed anti-voices effects. A mixed between within ANOVA investigated the effects of psychological inflexibility on the trial-types, but no effects were found (all p 's >0.05). Two independent t-tests showed no differences between the groups, and four one sample t-tests showed no significant differences for any trial-type from zero (all p 's > 0.05).

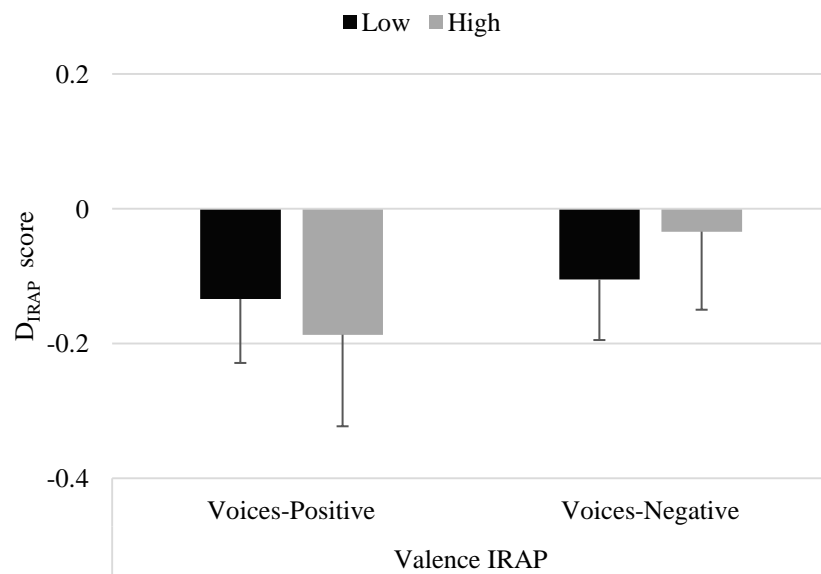


Figure 30. Mean D_{IRAP} scores for the two AAQ-II inflexibility levels on the Valence IRAP in Experiment 8. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

Perceived controllability analyses. The mean D_{IRAP} scores for the four perceived controllability dimensions as measured by the AHRS (i.e. full control [$N=11$], some-majority of time [$N=8$], some-occasionally [$N=7$], and no control [$N=7$]) on the Valence IRAP are presented in Figure 31. On Voices-Positive, three

groups (excluding some-majority) showed anti-voices effects, with the largest effect recorded in the some-occasional group, followed by no control and finally full control. The some-majority group showed null effects. On Voices-Negative, the same three groups (except some-majority) again showed comparable anti-voices effects, while the some-majority group showed marginal pro-voices effects.

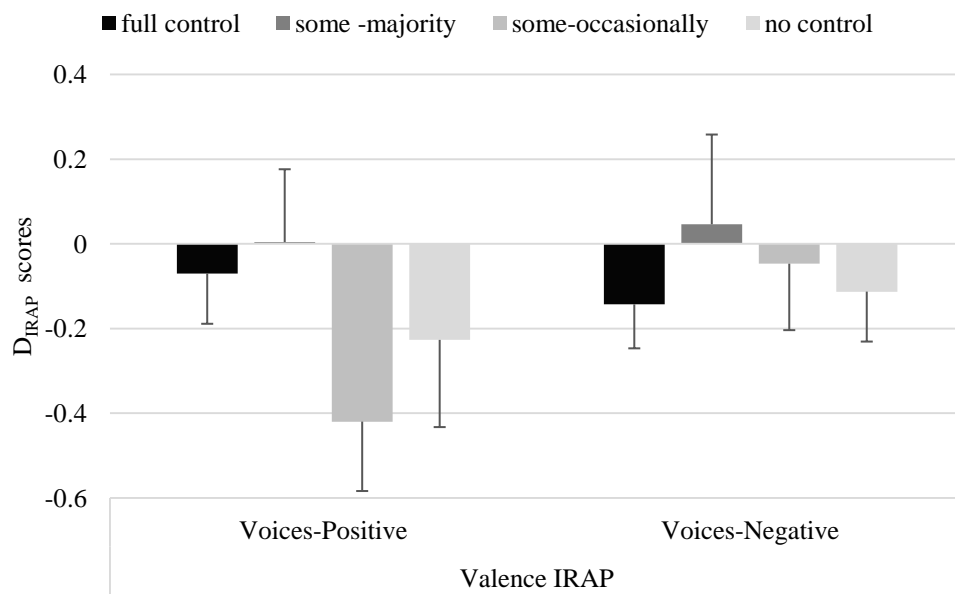


Figure 31. Mean D_{IRAP} scores for AHRS perceived controllability on the Valence IRAP in Experiment 8. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

A mixed between within ANOVA investigated potential differences among the perceived controllability groups and IRAP effects, but no effects were found (all $p's > 0.05$). Twelve independent t-tests, investigated group differences on each trial-type, but no differences were found (all $p's > 0.05$). Eight one-sample t-tests investigated whether the D_{IRAP} effects differed significantly from zero. Significant effects were found only recorded with some-occasionally on Voices-Positive ($df=6$, $t=-2.566$, $p < 0.05$).

Correlations. A correlation matrix also investigated potential relationships between the IRAP effects and the explicit measures. The only correlations recorded were between intensity of distress (on the AHRS) and Voices-Positive ($r=0.376$, $p<0.05$). That is, the more participants rated voices as positive, the greater their intensity of distress. For Voices-Negative, there were correlations with many AHRS subscales: duration ($r=0.348$, $p<0.05$); amount of negative content ($r=0.402$, $p<0.05$); degree of negative content ($r=0.353$, $p<0.05$); and intensity of distress ($r=0.402$, $p<0.05$). That is, the more participants confirmed voices as negative: the less duration of their voices; amount of negative content; degree of negative content and intensity of their distress.

Bonferroni adjustment. Due to multiple correlational analyses conducted, a Bonferroni adjustment of the alpha level was applied ($p<0.0005$). However, using this p-value, the above correlations do not reach significance (all p 's >0.0005). Indeed, while this adjustment is deemed necessary for statistical conservatism and to avoid the occurrence of a Type 1 error, for the purposes of the current thesis, the correlations which were significant at the 0.05 level will also be discussed.

Acceptance IRAP Data

Between groups analyses (clinical, non-clinical and controls). The mean D_{IRAP} scores for the three groups on the Acceptance IRAP are presented in Figure 32. On the Pleasant-Accept trial-type, the clinical and non-clinical groups showed acceptance effects (i.e. on average participants were faster responding pleasant voices-accept-true), whereas the control groups showed avoidance effects. On the Pleasant-Avoid trial-type, all three groups showed acceptance effects. On the Annoying-Accept trial-type, all groups showed marginal acceptance effects, however

the clinical group's effect was almost zero. On the Annoying-Avoid trial-type, both voice hearing groups showed avoidance effects, the largest of which were found in the clinical group, however, the non-voice hearers showed acceptance effects.

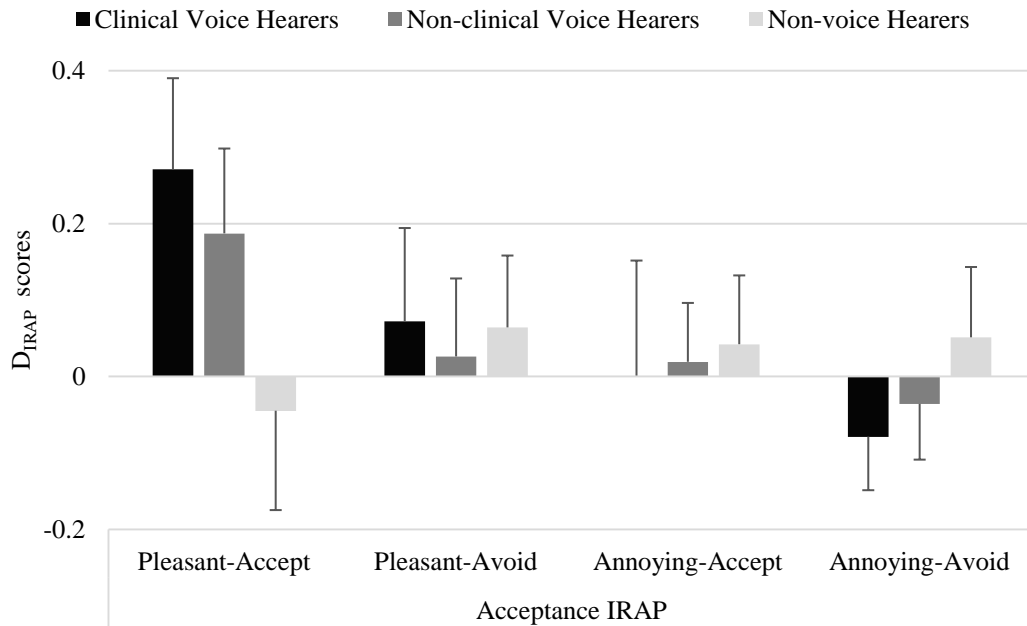


Figure 32. Mean D_{IRAP} scores for the three groups on the Acceptance IRAP in Experiment 8. **Positive D_{IRAP} scores indicate acceptance effects and negative D_{IRAP} scores indicate avoidance effects.**

To investigate the effects of group on trial-type on IRAP effects, a mixed between within ANOVA found no main or interaction effects (all p 's >0.05). Exploratory trial-type analyses as 12 independent t-tests, determined potential differences among the groups on individual trial-types, but again there were no significant differences (all p 's >0.05). Twelve one sample t-tests investigated whether the D_{IRAP} trial-type differed significantly from zero, but none did (all p 's >0.05).

Distress level analyses. The same distress level analyses have been applied for the Acceptance IRAP as the Valence IRAP using the DASS and the AAQ-II.

DASS. The mean D_{IRAP} scores for the three distress level on the Acceptance IRAP are presented in Figure 33. On Pleasant-Accept, all groups showed acceptance effects, the largest of which were found in the mildly distressed group. On Pleasant-Avoid, the mildly distressed group also showed acceptance effects, whereas the moderate and severe groups showed avoidance effects. On Annoying-Accept, the mild and severe groups showed marginal acceptance effects, whereas the moderate group showed avoidance effects. On Annoying-Avoid, the moderately distressed group showed marginal acceptance effects, whereas the mild (albeit very marginally) and severe group showed avoidance effects.

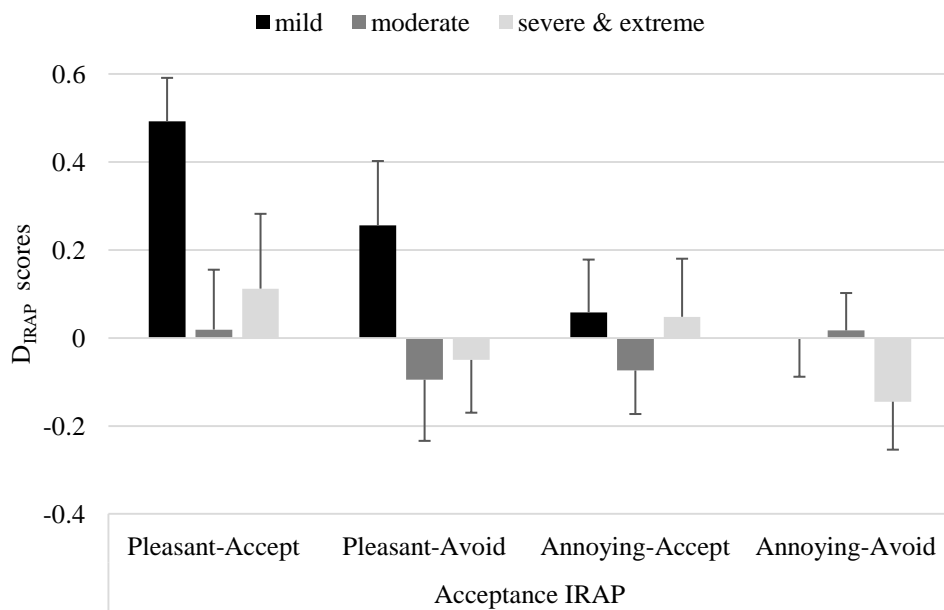


Figure 33. Mean D_{IRAP} scores for the three DASS distress groups on the Acceptance IRAP in Experiment 8. **Positive D_{IRAP} scores indicate acceptance effects and negative D_{IRAP} scores indicate avoidance effects.**

To investigate the effects of distress on trial-type, a mixed between within ANOVA found a significant main effect for distress ($F=3.644, p<0.05, \eta^2=0.623$). Post-hoc analyses as 12 independent t-tests found a significant difference between

the mild and moderate distress groups on Pleasant-Accept ($df=19, t=2.847, p<0.01$). Again, twelve one-sample t-tests investigated whether each of the D_{IRAP} effects differed significantly from zero. On Pleasant-Accept, the mildly distressed group was significant ($df=10, t=4.980, p<0.001$).

AAQ-II. The mean D_{IRAP} scores for each inflexibility level on the Acceptance IRAP are presented in Figure 34. On Pleasant-Accept, both groups showed acceptance effects, however for the low inflexibility group this was very marginal. On Pleasant-Avoid, the low inflexibility group showed avoidance effects, whereas the high inflexibility group showed acceptance effects. On Annoying-Accept, the low inflexibility group showed acceptance effects, whereas the high inflexibility group showed marginal avoidance effects. On Annoying-Avoid, both groups showed marginal avoidance effects, the larger of which was in the high inflexibility group.

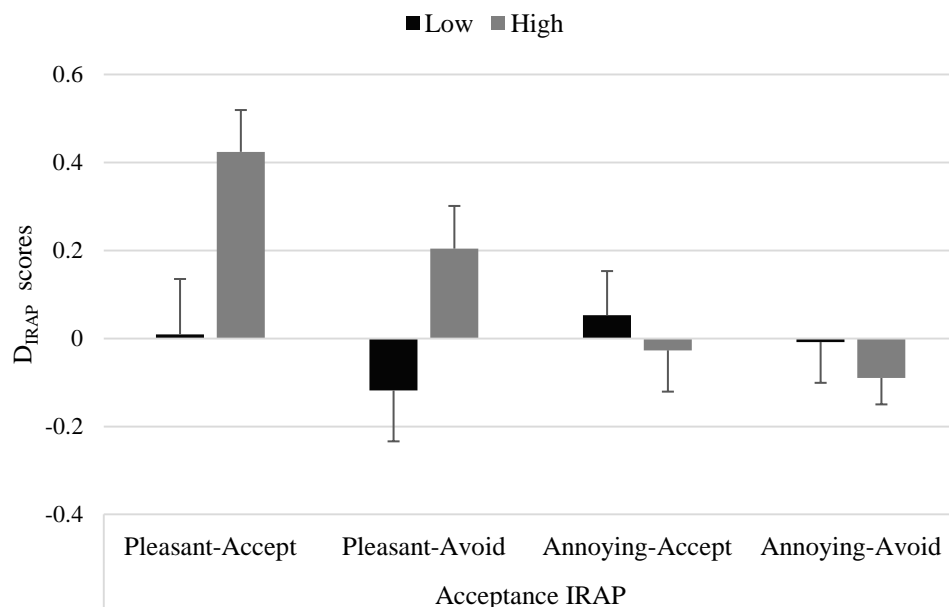


Figure 34. Mean D_{IRAP} scores for the two AAQ-II inflexibility levels on the Acceptance IRAP in Experiment 8. **Positive D_{IRAP} scores indicate acceptance effects and negative D_{IRAP} scores indicate avoidance effects.**

A mixed between within ANOVA demonstrated a significant main effect for trial-type ($F=2.878$, $p<0.05$, $\eta^2=0.668$) and an interaction effect ($F=3.836$, $p<0.05$, $\eta^2=0.811$). Post-hoc analyses as eight independent t-tests revealed a significant difference between the two inflexibility groups on Pleasant-Accept ($df=31$, $t=-2.603$, $p<0.05$) and Pleasant-Avoid ($df=31$, $t=-2.115$, $p<0.05$). Twelve dependent t-tests investigated the differences between levels of psychological inflexibility on each trial-type. No differences were found in the low group (all p 's >0.05). Within the high group, differences were found between Pleasant-Accept and Annoying-Accept ($df=15$, $t=3.671$, $p<0.01$), between Pleasant-Accept and Annoying-Avoid ($df=15$, $t=4.351$, $p<0.001$), and between Pleasant-Avoid and Annoying-Avoid ($df=15$, $t=2.518$, $p<0.05$). Again, eight one-sample t-tests found no significant effects in the low group (all p 's >0.05), but Pleasant-Accept was significant for the high group ($df=15$, $t=4.457$, $p<0.001$).

Perceived controllability analyses. The mean D_{IRAP} scores for the four perceived controllability groups on the Acceptance IRAP are presented in Figure 35. On Pleasant-Accept, three groups showed acceptance effects, the largest of which was in the some-majority group, however, the full control group showed marginal avoidance effects. Similarly, on Pleasant-Avoid, three groups showed acceptance effects, except the full control group who showed avoidance effects. On Annoying-Accept, the full control and some-majority groups showed avoidance effects, while some-occasionally and no control groups (albeit very marginal) showed acceptance effects. On Annoying-Avoid, the full control and no control groups showed marginal acceptance effects, while the some-majority and some-occasionally groups (albeit very marginal) showed avoidance effects.

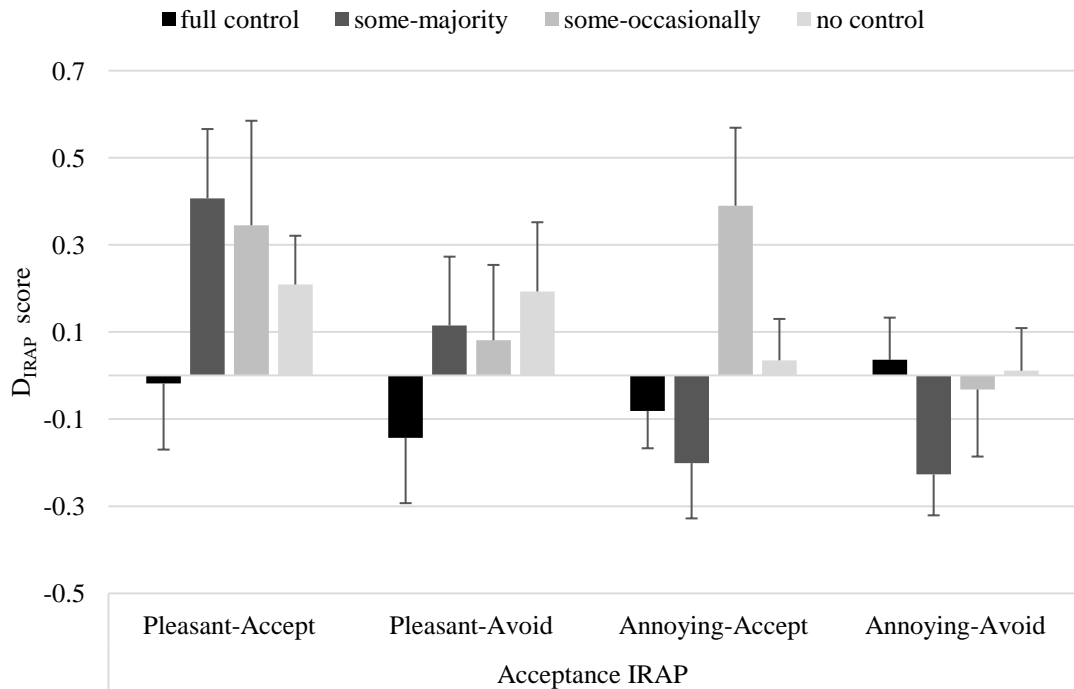


Figure 35. Mean D_{IRAP} scores for AHRS perceived controllability on the Acceptance IRAP in Experiment 8. **Positive D_{IRAP} scores indicate acceptance effects and negative D_{IRAP} scores indicate avoidance effects.**

A mixed between within ANOVA demonstrated a significant main effect for trial-type ($F=3.068$, $p<0.05$, $\mu_2 =0.699$). Post-hoc analyses, as 24 independent t-tests, investigated group differences on each trial-type. On Annoying-Accept, full control and some-occasional control groups again differed significantly ($df=16$, $t=-2.650$, $p<0.05$), as did some-majority and some-occasional ($df=13$, $t=-2.748$, $p<0.05$). Sixteen one-sample t-tests found significant effects for the some-majority group on Pleasant-Accept ($df=7$, $t=2.561$, $p<0.05$) and Annoying-Avoid ($df=7$, $t=-2.423$, $p<0.05$).

Correlations. A correlation matrix also investigated potential relationships between IRAP effects and explicit measures. Significant correlations are presented in Table 16. For Pleasant-Accept, there was a positive correlation with the AAQ-II, where the more participants accepted pleasant voices, the greater their inflexibility.

There were a number of negative correlations between this trial-type and various features of the explicit measures including: malevolent appraisals (BAVQ-R); omnipotent appraisals (BAVQ-R); depressive frequency (CAPE); depressive distress (CAPE); negative frequency (CAPE); negative distress (CAPE); total DASS; depression (DASS); anxiety (DASS); stress (DASS); and loudness of voices (AHRs). That is, the less participants accepted pleasant voices, the greater: malevolent and omnipotent appraisals; frequency of depressive symptoms and distress associated with their depressive symptoms; negative frequency of symptoms and distress associated with their negative symptoms; overall distress; and loudness of voices. For Pleasant-Avoid, there was a positive correlation with the AAQ-II, where the more participants avoided pleasant voices, the greater their inflexibility. There were also negative correlations with omnipotence appraisals (BAVQ-R) and depression (DASS). That is, the more participants avoided pleasant voices, the greater their appraisals of omnipotence and depression. For Annoying-Accept, there was a positive correlation with the VAAS, where the more participants accepted annoying voices, the greater overall acceptance of voices. For Annoying-Avoid, there was a negative correlation with frequency of voices, where participants avoided annoying voices, the greater their frequency of voices.

Table 16

Correlation table outlining significant correlations between Acceptance IRAP trial-types and the explicit measures in Experiment 8

	AAQ -II	Malev. (BAVQ)	Omnip. (BAVQ)	DASS	VAAS	Dep. Freq. (CAPE)	Dep. Dist. (CAPE)	Neg. Freq. (CAPE)	Neg. Dist. (CAPE)	Dep. (DASS)	Anx. (DASS)	Stress (DASS)	Freq. of Voices (AHRs)	Loudness (AHRs)
Pleasant- Accept	.375*	-.360*	-.458**	-.408*	-.100	-.345*	-.345*	-.407*	-.452**	-.387*	-.403*	-.360*	-.209	-.349*
Pleasant- Avoid	.389*	-.247	-.397*	-.323	-.189	-.057	-.075	.013	-.049	-.398*	-.240	-.352	-.234	-.288
Annoying -Accept	-.012	-.103	-.168	-.070	.358*	-.087	-.122	-.072	-.064	-.136	.031	-.083	.160	.135
Annoying -Avoid	-.138	-.180	-.158	-.072	-.309	.138	.032	.112	0.33	-.180	-.016	-.020	-.367*	-.002

* $p < 0.05$, ** $p < 0.01$

Bonferroni adjustment. Due to multiple correlational analyses conducted, a Bonferroni adjustment of the alpha level was applied ($p < 0.0005$). However, using this p-value, the above correlations do not reach significance (all p 's > 0.0005). Indeed, while this adjustment is deemed necessary for statistical conservatism and to avoid the occurrence of a Type 1 error, for the purposes of the current thesis, the correlations which were significant at the 0.05 level will also be discussed.

Summary and Discussion

As expected, controls evaluated voice hearing positively on the Valence IRAP, but for the first time, non-clinical voice hearers showed implicit negativity, and we had now recorded for the second time, negativity by the clinical voice hearers. The distress level analyses showed voice hearers who were moderately distressed, who were high on inflexibility, or who reported occasional or no control, were the most negative, which complements previous findings that negative beliefs and appraisals about voices predict distress (Peters et al., 2012; van der Gaag et al., 2003). This negativity correlated with: longer duration of voices; more negative content; and greater intensity of voice-related distress.

The Acceptance IRAP allowed us to juxtapose pleasant and annoying voices and the two voice hearing groups were similar in both respects. That is, clinical voice hearers had greater implicit acceptance of pleasant voices and avoidance of annoying voices than non-clinical voice hearers, which complements previous reports that the latter group use less avoidance and maladaptive coping strategies (e.g. Brett et al., 2007). Indeed, the clinical voice hearers' acceptance of pleasant voices was significant. Again, these effects were influenced by other factors. Specifically, the mildly distressed showed greatest acceptance of pleasant voices, while the severely

distressed showed most avoidance of annoying voices, which concords with previous findings that malevolent appraisals is related to voice resistance (e.g. Peters et al., 2012). The highly avoidant showed significantly more acceptance of pleasant voices, and avoidance of annoying voices. Those with some control showed the most significant acceptance of pleasant voices and avoidance of annoying voices. Among the various correlations, one of the most interesting indicated that acceptance of pleasant voices correlated with high psychological inflexibility, and acceptance of annoying voices correlated with overall voice acceptance.

Experiment 8 interestingly showed greater overall similarities between the two groups of voice hearers than was observed in the previous study, and for the first time non-clinical voice hearers were implicitly negative about voices. This suggested that the Valence IRAP was particularly suited to separating out voice hearers from controls. However, it was the Acceptance IRAP and its juxtaposition of pleasant and unpleasant voices which was the key feature of Experiment 8. And indeed, the findings were highly intuitive when the results showed that voice hearers were more accepting of pleasant voices and more avoidant of annoying voices. This was precisely the level of psychological precision that we had been working towards throughout the thesis and the refinements and the systematic manipulations of the IRAP have led us steadily towards this point from simple assessments of valence by control participants to distinguishing between the different types of reactions that different groups of voice hearers might have toward different types of voices. Once again, these effects supported those we had observed previously, but critically showed that the clinical voice hearers were implicitly less positive than the other two groups.

Limitations

There are a number of limitations of Experiment 8 which should be reflected on to guide future research. 1. The time-point at which participants were hearing voices (i.e. past/present) was not controlled for and may have influenced the distress-level analyses. 2. The English-to-Dutch translation of IRAP stimuli was conducted by a team of Dutch researchers and clinicians, however, this was not conducted using a blind backward-forward method which is the gold-standard for translation in research studies. This is a potential confounding variable in this research. 3. The use of inferential statistics in low N analyses (i.e. the distress level analyses) does not allow researchers to observe the differences or effects which may exist. Future research should try to include larger N 's for the distress level analyses in order to better examine this nuanced relationship between one's own level of distress and implicit attitudes.

Chapter 9
*Systematic comparisons across
experiments 1-8*

Systematic Experimental Comparisons

Each of the eight preceding experiments sought to assess implicit attitudes toward psychological suffering, especially hearing voices, in control, non-clinical and clinical samples ($N=377$). Each experiment investigated various aspects of these phenomena including: evaluations and stigmatisation of suffering; locus of control of suffering; evaluations of voices; perceptions of voices as normal; fear of voices; self-related evaluations of voices; other-related evaluations of voices; positive or negative valence of voices; and acceptance of positively and negatively perceived voices.

There were strong overlaps in the experimental designs and analytic strategies employed across the studies, which permitted useful comparisons of the response patterns of the various groups and of the range of psychological phenomena being investigated with the implicit and explicit measures. The current chapter is designed to review the studies and their findings, and to draw out the most interesting comparisons across these. Hence, the chapter is divided into three core sections. In Section 1, the experimental design and analytic strategy adopted in each study is summarised. In Section 2, the findings from each study are summarised and reviewed. And finally, in Section 3, systematic comparisons are drawn across the datasets, in terms of between groups analyses, distress level analyses and the predictive validity of the IRAP.

Section 1: Experimental Designs and Analytic Strategies

Experiments 1-3: Assessing stigmatisation and locus of control for psychological suffering, as well as evaluations of voice hearing in non-clinical samples. The experimental designs and analytic strategies adopted in Experiments 1-

3 were identical. All three studies involved non-voice hearing participants completing one IRAP each. The six IRAPs employed across the studies had broadly similar trial-types that juxtaposed positive versus negative evaluations of various aspects of psychological suffering. Analyses of variance then explored how performances on the trial-types varied across the IRAPs. Experiments 1 and 2 also involved distress level analyses of depression, anxiety and stress on the DASS to investigate the potential influence of this distress on the IRAP effects as observed. Because Experiment 3 employed the first IRAP on voice hearing, the distress level analyses there were based on psychotic-like experiences, as measured on the CAPE. Hence, these analyses explored the potential influence of these experiences on the effects on the voice hearing IRAP. Experiments 1-3 also employed correlational analyses to investigate the predictive validity of these IRAPs.

Experiment 4: Assessing the impact of a voice hearing simulation on voice hearing IRAPs in non-voice hearing participants. Experiment 4 differed in several ways from the three previous studies. Although it similarly involved non-voice hearing participants presented with a single IRAP that largely resembled the voice hearing IRAP from Experiment 3, this study employed a pre-post experimental design to determine the impact of a voice hearing simulation on the IRAP. Analyses of variance then examined the impact of the simulation on the IRAP effects (pre vs. post) and distress level analyses on the CAPE again explored the influence of psychotic-like experiences on the voice hearing IRAPs. Again, correlational analyses investigated the predictive validity of this voice hearing IRAP.

Experiments 5 and 6: Assessing the normality of voice hearing, as well as the evaluation of self and others hearing voices in non-clinical voice hearers and controls. Experiments 5 and 6 were the first to conduct between groups analyses to

compare control participants with non-clinical voice hearers. However, Experiment 5 involved distress level analyses on the CAPE to explore the influence of psychotic-like experiences on the Normality IRAP, while Experiment 6 involved distress level analyses on the AAQ-II to explore the influence of psychological inflexibility on the Self and Others IRAPs. Both studies, again, used correlational analyses to investigate the predictive validity of the IRAPs.

Experiments 7 and 8: Assessing the fear, valence and acceptance of voices in clinical and non-clinical voice hearers and controls Experiments 7 and 8 also conducted between groups analyses but did so to compare controls, non-clinical voice hearers *and clinical voice hearers*. The variations in these three IRAPs called for variations in the explicit data that comprised the distress level analyses. While both studies explored the influence of psychological inflexibility (AAQ-II) and perceived controllability (AHRs) on the Fear, Valence and Acceptance IRAPs, Experiment 7 also examined the influence of psychotic-like experiences (CAPE), whereas Experiment 8 examined the influence of distress (DASS). Again, both studies employed correlational analyses to investigate the predictive validity of these IRAPs.

Taken together, the consistency in the experimental designs and analytic strategies generated a wealth of information from multiple samples that would nonetheless permit some level of systematic comparison primarily across participant groups, across IRAPs, and across explicit and implicit measures. In order to draw systematic comparisons of the data across the studies, the following section summarises the findings from each study.

Section 2: Summaries of Study Findings

Experiment 1. This study assessed evaluations of (and stigmatisation to) psychological suffering, by juxtaposing suffering with normality using positive versus negative comparisons. Each of three IRAPs targeted a specific type of suffering, namely depression (using the Depression IRAP), anxiety (using the Anxiety IRAP) and the concept of mental illness generally (using the Mental Illness IRAP). Overall, all three IRAPs produced similar patterns of responding, which, surprisingly, indicated that all three forms of suffering were implicitly evaluated as **both** positive (e.g. anxious people are-relaxed-true) *and* negative (e.g. depressed people are-lazy-true). Similarly, all three IRAPs also showed pro-normality effects (e.g. normal people are-happy-true), although an anti-normality effect (e.g. normal people are-sad-true) was found on the Mental Illness IRAP.

In spite of the consistency of implicit effects across the three suffering-based IRAPs in Experiment 1, the distress level analyses generated highly divergent effects. On the Depression IRAP, participants who explicit as *mildly* distressed (on the DASS) implicitly evaluated depression most negatively, while the *severely* distressed implicitly evaluated normality most positively. However, the effects were somewhat different for the Anxiety IRAP, on which the *severely* (rather than mildly) distressed implicitly evaluated anxiety most negatively and also evaluated normality negatively. On this IRAP, the *mildly* and *moderately* (rather than severely) distressed evaluated normality most positively. On the Mental Illness IRAP, a similar level of implicit negativity to mental illness was recorded across all distress groups, but only the mildly or severely distressed also showed positive mental illness evaluations. The groups also showed a similar mix of positive and negative implicit evaluations of normality. Taken together, these data suggest that there is a relationship between

psychological suffering and evaluations of suffering and normality. However, the data also show that these effects are subtle and complex (e.g. the most severely distressed were the most stigmatising towards anxiety).

The correlations on the data from Experiment 1 revealed that the IRAP trial-types were predictive of some aspects of explicit. For the Depression IRAP, positive implicit evaluations of depression correlated with low anxiety: positive and negative implicit evaluations correlated with low and high explicit stigma respectively; and positivity toward normality effects correlated with high distress. For the Anxiety IRAP, pro-normality effects again correlated with overall distress. Overall, the correlations demonstrated that psychological distress impacts upon stigmatisation.

Experiment 2. The two IRAPs employed in Experiment 2 assessed perceptions of the locus of control for, and evaluations of, psychological suffering by juxtaposing suffering with health using positive versus negative comparisons. Each IRAP targeted suffering as either an illness (using the Illness IRAP) or a weakness (using the Weakness IRAP). Both IRAPs produced similar patterns of responding, in which suffering (as illness or weakness) was implicitly evaluated as **both** positive (e.g. ill people are-happy-true) *and* negative (e.g. ill people are-depressed-true). Some strong positivity toward health was also recorded.

Again, however, the distress level analyses generated divergent effects. Similar to the Depression IRAP, the *mildly* distressed evaluated illness most negatively and health most positively on the Illness IRAP. This was also the case for the Weakness IRAP in which the *mildly* distressed evaluated weakness most negatively, although in this context it was the *severely* distressed who evaluated health most positively. Furthermore, the correlations revealed that positivity toward health correlated with high explicit stigma on the Illness IRAP. On the Weakness

IRAP, negativity toward weakness correlated with low distress, while positivity toward health correlated with more distress and stigma.

Experiment 3. Experiment 3 was the first to assess explicit and implicit attitudes toward hearing voices, using an Evaluation IRAP that juxtaposed hearing voices with seeing things using positive versus negative comparisons) and presented to non-voice hearing participants. Overall, hearing voices was implicitly positively evaluated. The distress level analyses showed that participants who explicit as *high* on psychotic-like experiences (on the CAPE) evaluated voices positively, while the *low* were even more positive, but also showed marginal negativity. Surprisingly, implicit positivity toward voice hearing correlated with high stigma, but again negativity correlated with high distress.

Experiment 4. This experiment sought to investigate the malleability of fearful evaluations of hearing voices, using a Fear IRAP and a voice hearing simulation. At baseline and post-simulation, hearing voices was implicitly evaluated as **both** positive *and* fearful, although positivity toward voices reduced and fear increased after the simulation. Distress level analyses indicated that participants *low* on psychotic-like experiences showed both positivity and fear toward voices at both pre- and post-simulation, with little change in between. However, for participants who were *high*, there was less positivity and more fear at post-simulation. Interestingly, the correlations between the IRAP and the explicit data did not yield any significant results.

Experiment 5. Experiment 5 was the first to involve non-clinical voice hearers and used the Normality IRAP to assess evaluations of hearing voices as *normal*. Both non-clinical voice hearers and non-voice hearing controls implicitly evaluated voice hearing as normal, although the effect was stronger for the voice

hearers. All voice hearers explicit as high on psychotic-like experiences according to the CAPE. Participants who were high in this regard evaluated voices as most normal, while the *low* evaluated voices as least normal and showed marginal anti-normality reactions to voice hearing. This effect for low CAPE participants matched Experiments 3 and 4. Evaluations of voice hearing as normal correlated with behavioural engagement with voices as well as voice acceptance, while evaluations of voice hearing as abnormal correlated with high benevolent appraisals and high emotional engagement with voices and other psychotic-like experiences.

Experiment 6. This experiment sought to investigate potential differences in voice hearing and control participants' evaluations of the *self* and *others* hearing voices, using the Self IRAP and the Others IRAP. Overall, hearing voices was evaluated positively by both groups on both IRAPs, although control participants were more positive on the Self IRAP. Interestingly, control participants were also more fearful on the Self IRAP, while the voice hearers were more fearful on the Others IRAP. That is, voices are more frightening and harder to accept if I have no experiences of hearing them, but if I have experience of hearing them, it would seem that they are more frightening and hard to accept for other people.

In spite of the consistency of implicit effects, the distress level analyses again generated divergent effects. Voice hearers who reported *high psychological inflexibility* on the AAQ-II were unsurprisingly less implicitly fearful of voices, whereas the low in inflexibility were more fearful on the Self and Others IRAPs. Correlations revealed that, for the voice hearers, implicit positivity to voices on the Self IRAP correlated with high depressive CAPE symptoms and implicit fear on the Others IRAP correlated with higher voice acceptance.

Experiment 7. This was the first study to involve clinical voice hearers and compared these with non-voice hearing controls and non-clinical voice hearers on the Fear IRAP. Similar to Experiment 4, non-clinical voice hearers and controls implicitly evaluated hearing voices as **both** positive *and* fearful, but clinical voice hearers showed implicit negativity. As in Experiments 3 and 4, all voice hearers who were *low* on psychotic-like experiences showed positivity to voices, while those *high* on psychotic-like experiences *and high* on psychological inflexibility evaluated voices as most fearful. When voice hearers were divided according to level of perceived control over their voices (on the AHRS), highly divergent effects were observed. Voice hearers who reported full control evaluated voices most positively, while those with some or no control evaluated voices as most fearful. The correlations showed that implicit fear of voices correlated with greater voice severity (AHRS total score), while positivity correlated with behavioural resistance to voices.

Experiment 8. This experiment again sought to explore differences among clinical, non-clinical voice hearing and non-voice hearing control participants' in terms of their emotional (using the Valence IRAP) and behavioural responses (using the Acceptance IRAP) to hearing voices.

While controls evaluated voice hearing positively on the Valence IRAP, as they had done in Experiments 3-6, voice hearing was evaluated negatively by both groups of voice hearers. This effect for the clinical voice hearers was similar to Experiment 7. Voice hearers who were *moderately* distressed evaluated voices most negatively, as did those who were *high* on psychological inflexibility. Dividing voice hearers according to perceived control over voices, however, created some divergence. Voice hearers who reported occasional or no control evaluated hearing voices most negatively, as in Experiment 7. And the correlational data found that this

negativity correlated with: lower duration of voices; less negative content; and less intensity of voice-related distress.

On the Acceptance IRAP, clinical and non-clinical voice hearers implicitly accepted positive voices and avoided negative voices. Those with *mild* distress showed greatest implicit acceptance of positive voices, while the *severely* distressed showed the *largest* implicit avoidance of negative voices. Indeed, voice hearers who were *highly* avoidant on the AAQ-II showed greatest implicit acceptance of positive voices *and* avoidance of negative voices. Similarly, voice hearers with some perceived control showed greatest implicit acceptance of positive voices and avoidance of negative voices. The correlations showed that implicit acceptance of positive voices correlated with high psychological inflexibility, and low acceptance of positive voices correlated with higher: malevolent and omnipotent appraisals; depressive symptoms; negative symptoms; overall distress; and loudness of voices. Furthermore, avoidance of positive voices correlated with omnipotent appraisals and depression. And interestingly, acceptance of negative voices correlated with overall voice acceptance and avoidance of negative voices correlated with greater voice frequency.

Systematic Experimental Comparisons

Between group analyses. A number of key patterns in the data emerged from the between group analyses and these are summarised below:

- Participants showed both implicit positivity and negativity (stigmatising and non-stigmatising) to various labels of psychological suffering, including anxiety, depression and the concept of mental illness generally.

- When the IRAPs presented voice hearing specifically, non-voice hearing participants implicitly evaluated this experience as positive and normal, although they did show implicit fear.
- Interestingly, when presented with the voice simulation procedure, non-voice hearing participants showed less implicit positivity and more fear after direct experience of voices through the simulation.
- In general, the non-clinical voice hearers showed IRAPs effects that more often resembled those of the non-voice hearing control participants, rather than clinical voice hearers, although this varied across IRAPs.
- Similar to non-voice hearing controls, non-clinical voice hearers implicitly evaluated voice hearing as positive and normal, although they once again showed implicit fear and some negativity.
- Interestingly, this implicit positivity toward voice hearing when the IRAP co-ordinated voice hearing with the self and with others, but some negativity was recorded in the context of voices heard by others.
- Non-clinical voice hearers also showed implicit acceptance for both positive and negative voices, and implicit avoidance for negative voices.
- In contrast to both non-voice hearing controls and non-clinical voice hearers, clinical voice hearers showed implicit negativity and fear of voice hearing, and not surprisingly therefore, they also showed implicit acceptance of positive voices and avoidance of negative voices.

Distress level analyses. A number of key patterns in the data emerged from the distress level analyses and these are summarised below.

- While participants who were *mildly* distressed on the DASS were most implicitly negative (stigmatising) about the concepts of depression, illness

and weak-mindedness, *severely* distressed participants were most implicitly negative about *anxiety*.

- In accordance with these effects, the *mildly* distressed also implicitly evaluated normality as positive when juxtaposed with anxiety and illness, while the *severely* distressed evaluated normality as positive when juxtaposed with depression and weak-mindedness.
- Participants with *low* psychotic-like experiences (on the CAPE) implicitly evaluated voice hearing as positive and normal, while also showing implicit negativity, fearfulness and evaluations of abnormality. Similar patterns were recorded for those with *high* psychotic-like experiences, except that these were less fearful overall.
- While the voice hearing simulation had little impact on the low CAPE group, the simulation appeared to increase implicit negativity of voice hearing for the high group.
- As expected, both groups of voice hearers produced higher CAPE scores overall, but those who were *low* in psychotic-like experiences scores showed implicit positivity and fear of voices. This contrasted with little or no effects recorded with the voice hearers.
- Somewhat divergent patterns were also recorded when voice hearers were sub-divided by level of distress (on the DASS). That is, while voice hearers with *moderate distress* showed the greatest implicit negativity to voices, the *mild* showed greatest implicit acceptance of positive voices and the *severe* showed implicit avoidance of negative voices.
- More strongly divergent patterns were also recorded when voice hearers were sub-divided by level of psychological inflexibility (on the AAQ-II). Those

low on inflexibility showed both implicit positivity and fear of voice hearing when associated with self, but only positivity with others' voices. In contrast the *high* group showed positivity to voices when associated with self, but positivity and fear of others hearing voices. This latter group also showed implicit avoidance of negative voices.

- Strongly divergent patterns were also recorded when voice hearers were subdivided by level of perceived control over their voices (on the AHRS). Voices hearers with a sense of *full control* showed implicit positivity to voices. In contrast, those with *some control* showed implicit negativity and fear, while also showing acceptance of positive voices and avoidance of negative voices. And, unsurprisingly, those with *no control* showed implicit negativity toward voice hearing.

Correlational analyses. A number of key patterns emerged from the correlational analyses between the IRAP trial-types and explicit measures, and these are summarised below.

- High negativity across multiple IRAPs correlated with higher overall distress, and interestingly positivity toward normality and health also correlated with higher distress and greater stigma.
- For voice hearers, implicit negativity toward voices correlated with less voice severity on the AHRS (i.e. duration, amount of negative content and degree of negative content). And implicit positivity toward voices correlated with greater intensity of distress.
- Implicit evaluations of voice hearing as normal correlated with behavioural engagement with, and acceptance of, voices, while other aspects of voice positivity correlated with behavioural resistance.

- Implicit evaluations of voice hearing as abnormal correlated with more psychotic-like experiences, greater emotional engagement with voices and a greater appraisal of voices as benevolent. Implicit *fear* of voices correlated with less voice severity on the AHRS.
- Implicit positivity to voices in the context of self, correlated with more depressive psychotic-like symptoms, while fear of others' voices correlated with voice acceptance.
- *High* acceptance of positive voices correlated with high psychological inflexibility, whereas *low* acceptance correlated with greater voice loudness, malevolent and omnipotent appraisals, as well as greater depressive psychotic-like symptoms, negative psychotic-like symptoms and overall distress. High avoidance of positive voices also correlated with omnipotent appraisals and depression.
- Implicit acceptance of negative voices correlated with overall voice acceptance, while avoidance of negative voices correlated with high voice frequency.

Conclusions

The aim of the current chapter was to generate a broad set of conclusions regarding the findings of the current thesis. It is inherently difficult to draw conclusions from any individual study, or even across studies, but the current uniform analytic strategy we adopted permitted the best possible conclusions across even diverse IRAPs, and allowed systematic comparison of very specific experimental outcomes. Consider, therefore, what those broad conclusions are. 1. Participants overall showed implicit positivity towards concepts pertaining to

psychological suffering. 2. In numerous respects, non-clinical voice hearers responded in ways that were more similar to non-voice hearing controls than to clinical voice hearers. 3. Participants' explicit level of distress, psychological inflexibility and perceived control over voices appear to influence their IRAP effects, thus suggesting a complex interplay between these variables. Having drawn these conclusions, we hoped to set the scene for the next chapter, where we look at the points of contact and or discord between the findings here and the relevant literatures, particularly those on stigma to psychological suffering, psychosis and voice hearing. This was the aim of the general discussion in the final chapter -- Chapter 10.

Chapter 10
General discussion

General Discussion

IRAP research is defined largely by the presence of IRAP effects, but the precise nature and size of these vary considerably across studies. In short, IRAP effects are recorded as any effect that differs from a zero D_{IRAP} score, but at a more stringent technical level, one could argue that an IRAP effect can only be recorded if this D_{IRAP} score differs *significantly* from zero. However, many IRAP studies do not hold strictly to the latter because the N s involved in the studies, especially for example in clinical IRAP studies, are often relatively small, especially given the types of statistical analyses that are typically employed. Indeed, it could be argued that this method impacts upon on the credibility of IRAP research, however, in a recent meta-analysis of IRAP studies, it was reported that even small N IRAP studies have sufficient statistical power (see Vahey, Nicholson, and Barnes-Holmes, 2015). In any case, IRAP researchers reflect upon the meaning of their data and analyses, while recognising that their interpretations must be somewhat cautious when non-statistical but interesting or predicted IRAP effects are recorded.

Although some of the effects in the current thesis were not statistically significant from zero, they nonetheless appeared to be *functionally* meaningful. Consider, for example, effects recorded on a Voices-Negative trial-type. A D_{IRAP} score of zero would indicate that participants responded to Voices-Negative-True and Voices-Negative-False at equal speed. For RFT, this suggests that according to an individual's verbal history, there is the same likelihood of deriving one relation over the other, demonstrating relational flexibility (or ambivalence) for either relation. On the other hand, a D_{IRAP} score which is statistically different from zero suggests a history of relating those stimuli more rigidly, relatively speaking. For

example, on Voices-Negative, if there was a D_{IRAP} score of -0.5, participants had responded to Voices-Negative-True more quickly than Voices-Negative-False, because of a history of more derivation of the former. The IRAP, thus points toward the key functional processes behind patterns of responding and may help when categorising behaviour as flexible or rigid, for example. Furthermore, it is worth noting that if there was anti-voices effect on the Voices-Positive trial-type, for example, this suggests (functionally speaking) that participants *rejected positivity* toward voices (i.e. participants responded more quickly on Voices-Positive-False), which is functionally different to saying there was an anti-voices effect on this trial-type. And we would argue that even those small functional distinctions may specify a key functional pattern of responding.

Stigma Research

The current findings, especially from Experiment 1, appear to complement and extend the existing literature on the stigmatisation of psychological suffering. Overall, there were no differences in implicit stigma toward depression, anxiety and mental illness, which contrast with findings, such as those of Lincoln et al. (2008). There was however, evidence of a rejection of stigma towards anxiety (i.e. participants responded more quickly to anxious people are-nervous-false), which may be due to the high prevalence of anxiety in the general population. And most interestingly, levels of personal distress appeared to influence IRAP effects, which concord with the findings by Rüsçh et al. (2010), in which self-stigma predicted a lower quality of life. In our study, participants with less psychological contact with suffering showed greater stigma, and those who had most psychological contact with suffering showed greater positivity toward normality.

In Experiment 2, stigmatisation was slightly greater (but not significantly) when externalised as illness as a source of locus of control, which contrasts with effects reported by Teachman et al. (2006), who found no stigmatising effects for physical illness. Indeed, our study actually recorded global positivity toward being 'healthy'. Again, when investigating stigma toward voice hearing in Experiment 3 for example, we found none, and this differs from the stigma toward schizophrenia reported by Peris and colleagues (2008). Thus, it appears that there may be greater stigmatisation toward the concept of 'schizophrenia', than toward the topographies (at least hearing voices) often associated with this form of suffering.

In the context of the hearing voices simulation, the positive effects found in the literature where there were reductions in stigma was not replicated in Experiment 4 here (Dearing & Steadman, 2009; Deegan, 1996; Kidd et al., 2015; Sideras et al., 2015; Wieland, et al., 2015; Wilson et al., 2009). Specifically, we found a reduction in positivity and an increase in fear, and this effect was greater for those more familiar with psychotic-like experiences. However, negative simulation outcomes have been recorded elsewhere (e.g. Brown et al., 2010; Kalyanaraman et al., 2010). Hence, a useful path for future research may involve investigating the relationship between individual's own experiences and distress and the malleability of their implicit attitudes to hearing voices or other psychotic-like experiences. Indeed, the current thesis found evidence of this complex relationship in almost every study we conducted.

Predictive validity of the IRAP in the context of stigma research. Implicit stigmatisation and pro-normality against depression correlated with higher levels of anxiety and depression, respectively. And in the context of anxiety, positivity toward normality interestingly correlated with higher levels of anxiety. So, levels of

depression and anxiety appear to influence stigmatisation. For example, an anxious person may be more stigmatising toward depressive topographies because they are unlike their own patterns of responding, and thus cannot make psychological contact with this way of operating. In the context of illness, positivity toward health correlated with explicit stigma. And in the context of weak-mindedness, implicit stigma correlated with less distress. Negativity toward voices correlated with distress and interestingly, positivity toward voices correlated with higher psychotic-like experiences in non-voice hearing control participants. Once again, these findings highlight the influence of participants' own histories and experiences, as well as indicating the predictive utility of the IRAP in tapping into this source of influence.

Implicit Appraisals of Voice Hearers

Contrary to some evidence suggesting that sufferers categorise voice hearing as abnormal (Corrigan, 2004; Mak et al., 2007), the current research found largely positive evaluations of voices by both non-voice hearing controls and non-clinical voice hearers. Indeed, less avoidant voice hearers showed negativity toward the self hearing voices, while high avoiders showed positivity, which is consistent with the literature on avoidance where avoidance facilitates escape from fear (Luciano et al., 2013). Indeed, this positivity toward voices was not observed in clinical voice hearers.

In terms of emotionally valenced responding to voices, all voice hearers (i.e. both clinical and non-clinical) showed negativity toward voices. Indeed, when this effect was split according to distress level, the moderately distressed group showed most negativity. This finding complements previous studies which have reported that negative beliefs and appraisals about voices predict distress (Peters et al., 2012; van

der Gaag et al., 2003). Our research also showed that those who perceived only some control over voices were most negative. Once again, these findings highlight the influence of participants' own experiences on their implicit attitudes.

In terms of behavioural responding to both positive and negative voices, clinical voice hearers showed greater implicit acceptance of positive voices and implicit avoidance of negative voices than non-clinical voice hearers. This finding is somewhat consistent with existing research which suggests that non-clinical voice hearers use less avoidance and maladaptive coping strategies (i.e. only accept positive voices) than clinical voice hearers (e.g. Brett et al., 2007). Indeed, those with mild distress showed the greatest acceptance of positive voices, while those with severe distress showed the greatest avoidance of negative voices. This latter finding is consistent with the literature suggesting that there is an association between negative (or malevolent) appraisals and voice resistance (e.g. Peters et al., 2012). And interestingly, the voice hearers who were high on avoidance showed the greatest acceptance of positive voices. And those with no control showed flexibility around responding to negative voices, indicating a greater acceptance of their presence.

Predictive validity of the IRAP in the context of voice hearing. Implicit negativity to voices (in the context of abnormality) correlated with voice engagement, benevolent appraisals, and distress, which is largely consistent with the literature (Chadwick & Birchwood, 1994). Furthermore, less implicit fear of voices correlated with greater total voice severity (on the AHRS), perhaps indicating greater exposure and acceptance. However, this conclusion would contradict the finding that appraisals are independent of voice severity (e.g. van der Gaag et al., 2003). Implicit positivity to voices correlated with behavioural resistance and distress, while implicit

negativity correlated with less total voice severity (on the AHRS), and again the latter challenges the finding that appraisals are independent of voice severity (van der Gaag et al.). These findings suggest, at least, that researchers need to be cautious in how they interpret positivity, in that positivity may or may not reflect a ‘psychologically healthy’ coping strategy with voices.

Furthermore, less implicit positivity to the self hearing voices correlated with depressive psychotic-like symptoms, which complements previous evidence that there is an inverse relationship between benevolent appraisals and depression (van der Gaag et al., 2003), however, the current implicit research is not directly comparable to benevolent appraisals, participants may appraise a voice as benevolent but still produce a negative implicit effect. Furthermore, implicit negativity to others’ voices correlated with voice acceptance, which again, shows a clear splice between emotional and behavioural responses to voices, that is, you can negatively evaluate something *while still* showing acceptance towards it.

The best predictors of the explicit measures were found in the context of behavioural responding in Experiment 8. Specifically, implicit acceptance of *positive* voices correlated with higher avoidance, less malevolent appraisals, omnipotent appraisals, distress, psychotic-like experiences and voice loudness. Implicit avoidance of *positive* voices correlated with higher avoidance, depression and omnipotent appraisals, which appears to complement Peter’s et al. (2012) finding that omnipotent appraisals were related to resistance, and that this was also associated with distress. Implicit acceptance of *negative* voices correlated with voice acceptance, whereas greater implicit avoidance of *negative* voices correlated with higher voice frequency.

Implications for the Literature

Psychosis. The current programme of research sought to investigate implicit responding in voice hearing and non-voice hearing groups, and the potential role of distress in these reactions, rather than relying solely on explicit responding to understand the onset, maintenance and experiential nature of hearing voices. Specifically, this approach helped us to investigate very particular features of voice hearing, such as the relationship between acceptance/avoidance of voices and levels of distress. Overall, some of the findings were consistent with the literature on psychosis, while some were inconsistent. And yet, these findings do make a noteworthy contribution to the field. For example, we found that voice hearers were more positive about voices than one would typically expect, but this effect was clearly influenced by levels of distress, avoidance and perceived control. Furthermore, the voice hearers who were most avoidant of negative voices were most psychologically avoidant overall, most distressed and had only some level of perceived control over their voices. And critically, implicit acceptance of negative voices directly related to voice acceptance. This is the level of psychological precision needed to ask complex questions about this type of suffering, and the IRAP has allowed us to take the first steps towards a deeper understanding of the phenomenon of hearing voices. Furthermore, these findings speak directly towards the types of support or intervention that would be of most benefit to voice hearers, and critically those in distress. For example, the data suggest that facilitating voice acceptance over avoidance may offer therapeutic benefit.

Implicit measures. One of the primary aims of the current research was to investigate the potential utility of the IRAP in studying more complex patterns of human suffering, such as hearing voices. We wanted to take the first steps to

understanding this phenomenon by first taking a very broad approach to suffering, and once the utility of the IRAP was observed, we started to look at more complex behaviours which may be involved in voice hearing. And the results have clearly demonstrated that the IRAP delivered a high level of precision which could not have been achieved, we believed, using any other currently available implicit or explicit measures. For example, the IRAP was able to distinguish between the clinical and non-clinical groups by their patterns of responding in the context of fear and acceptance. Specifically, clinical voice hearers showed greater acceptance of positive voices and avoidance of negative voices than the non-clinical voice hearers. And when you splice up the effects by distress level, we see even higher levels of precision in this context. That is, those with milder distress are more likely to accept positive voices, whereas those who were severely distressed were more likely to avoid negative voices. And it is through this precision that we can begin to look at the functional processes at play in the voice hearing experience.

Functional analysis. As functional contextualists, we naturally began this research with a functional aim, which not only informs the basic science, but also clinical applications. So, being able to answer functional analytic questions about voice hearing as outlined here, has begun to allow us to identify the key processes involved in this suffering, which is our overarching aim. Thus, through this research, we can begin to move towards a functional analytic account of voice hearing in terms of maintenance and possibly even onset. This is done with the hope that these will allow us to better understand these experiences, and perhaps ultimately change them in the service of the individual.

Clinical Implications

Clinical psychology agrees that the use of functional analyses to guide therapeutic work provides greater precision, which in turn leads to an enhanced ability to predict and influence behaviour. In practice however, we are constrained by the currently available research methodologies with which we can conduct functional analyses to investigate group patterns of behaviour, however the IRAP has become an increasingly reliable and valid measure of this. Therefore, while the findings of the current thesis may be preliminary in nature, patterns of verbal behaviour were observed that may indeed inform therapeutic practices. For example, level of distress is clearly associated with implicit voice negativity and acceptance or avoidance of voices. Taken together, these findings may have clinical implications, particularly for clinical voice hearers. For example, some therapeutic approaches encourage voice acceptance, while others encourage distraction, particularly from negative voices. The current data suggest that acceptance of positive voices and avoidance of negative voices may form part of an individual's clinical distress regarding voice hearing, and thus the implication is that therapy might most usefully encourage acceptance of negative voices and less emphasis on accepting positive voices. In any case, the findings here suggest the potential benefits of the IRAP in permitting considerable precision in the assessment of one's evaluation of, and reactions to, the experience of hearing voices.

Limitations of the Research

While the current thesis has highlighted a number of interesting findings, there are a number of limitations. First, in the early studies, the IRAP stimuli used may not fully capture the negative connotations of stigma, and may simply reflect

the patterns of traits often observed in each form of suffering. Future studies could include stimuli which target stigmatising behaviours toward sufferers (e.g. I would stay away from these people, they are different to me etc.). Second, the concept of normality was selected in this study as a contrast category for suffering, however, in reality ‘normality’ (i.e. normal verbal behaviour) also encompasses aspects of suffering. Future studies could investigate the use of alternative contrast categories which adequately encapsulate the absence of specific forms of suffering or diagnoses. Third, many of the samples used comprised of a high proportion of psychology undergraduates exposed to some level of psychological training, which may account for some of the positivity, therefore, it would be interesting to replicate this study in a sample with no training in psychology. Fourth, the time-point at which participants were hearing voices (i.e. past/present) was not controlled for and may have influenced the distress-level analyses. Fifth, in Experiment 8, the different language administration between the Dutch and Irish site may have influenced IRAP effects due to the lack of blind forward/backward translation between English and Dutch. Sixth, the use of inferential statistics in low N analyses (i.e. the distress level analyses) does not allow researchers to observe the differences or effects which may exist. Future research should try to include larger N 's for the distress level analyses in order to better examine this nuanced relationship between one's own level of distress and implicit attitudes.

Recommendations for Future Research

For future research in the domain of psychosis using implicit measures, we would recommend the following three studies: 1. To investigate if a therapeutic intervention could increase acceptance of negative voices using a pre-post design

and the IRAP as an implicit measure of change. 2. To build on current psychosis models which purport that the sense of self is fractured in voice hearing, using the IRAP as an implicit measure. 3. To investigate if a 'self-based' RFT intervention could integrate the fractured aspects of the self in this type of suffering using a pre-post design and the IRAP as an implicit measure of change. These studies would provide a strong springboard from which an empirically-based functional analytic account of voice hearing and psychosis could be built.

Conclusions

The aim of this work was to understand voice hearing and psychosis in order to facilitate better care for sufferers. And although we are a long way off the better care which is needed, we want to contribute to a better understanding of this pattern of suffering. But for us, as functional contextualists, we would argue that you should not separate care from understanding, as understanding is an essential ingredient in better care. The current thesis, albeit somewhat abstract at times, can begin to speak to how understanding can help facilitate care. For example, finding that acceptance of negative voices is important and it relates to level of distress illustrates this important link. If this outcome alone can facilitate better care for these individuals, then what we have added to the understanding of hearing voices, albeit small and even tentative, has been worthwhile.

References

- Al-Issa, I. (1977). Social and cultural aspects of hallucinations. *Psychological Bulletin*, 84(3), 570-587.
- Al-Issa, I. (1995). The illusion of reality or the reality of illusion. Hallucinations and culture. *British Journal of Psychiatry*, 166(3), 368-73.
- Andrew, E. M., Gray, N. S., & Snowden, R. J. (2008). The relationship between trauma and beliefs about hearing voices: A study of psychiatric and non-psychiatric voice hearers. *Psychological Medicine*, 38(10), 1409-1417.
- Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the implicit relational assessment procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, 32(7), 169-177.
- Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *Psychological Record*, 60(3), 527–542.
- Barnes-Holmes, Y., Hussey, I., McEnteggart, C., Barnes-Holmes, D., & Foody, M. (in press). The relationship between relational frame theory and middle-level terms in acceptance and commitment therapy. In S. Hayes, D. Barnes-Holmes, R. Zettle, and A. Biglan (Eds.), *Handbook of contextual behavioral science*. New York: Wiley-Blackwell.
- Beavan, V., & Read, J. (2010). Hearing voices and listening to what they say: the importance of voice content in understanding and working with distressing voices. *Journal of Nervous and Mental Disease*, 198, 201–205.

- Beavan, V., Read, J., & Cartwright, C. (2011). The prevalence of voice-hearers in the general population: A literature review. *Journal of Mental Health, 20*(3), 281–292.
- Benkman, L. (1971). Locus of control and attitudes toward mental illness among mental health volunteers. *Journal of Consulting and Clinical Psychology, 38*(1), 84-89.
- Bentall, R. P. (2003). *Madness Explained: Psychosis and Human Nature*. London: Penguin Books.
- Bentall, R. P. (2004). *Madness Explained: Psychosis and Human Nature*. London, England: Penguin Books Ltd.
- Birchwood, M., & Chadwick, P. (1997). The omnipotence of voices: Testing the validity of a cognitive model. *Psychological Medicine, 27*, 1345-1353.
- Birchwood, M., Iqbal, Z., & Upthegrove, R. (2005). Psychological pathways to depression in schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience, 255*, 202–212.
- Blakemore, S. J., Smith, J., Steel, R., Johnstone, E. C., & Frith, C. D. (2000). The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: Evidence for a breakdown in self-monitoring. *Psychological Medicine, 30*(5), 1131-1139.
- Bond, F., Hayes, S., Baer, R., Carpenter, K., Guenole, N., Orcutt, H., ... Zettle, R. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*(4), 676–88.
- Boyle, M. (2004). Preventing a non-existent illness? Some issues in the prevention of "schizophrenia". *Journal of Primary Prevention, 24*(4), 445-469.

- Brett, C. M. C., Peters, E. P., Johns, L. C., Tabraham, P. A., Valmaggia, L., & McGuire, P. K. (2007). The appraisals of anomalous experiences interview (AANEX): A multi-dimensional measure of psychological responses to anomalies associated with psychosis. *British Journal of Psychiatry* 191(51), S23–S30.
- British Psychological Society. (2009). *Code of Ethics and Conduct: Guidance published by the Ethics Committee of the British Psychological Society*. Accessed from: http://www.bps.org.uk/sites/default/files/documents/code_of_ethics_and_conduct.pdf.
- Brown, S. A., Evans, Y., Espenschade, K., & O'Connor, M. (2010). An examination of two brief stigma reduction strategies: filmed personal contact and hallucination simulations. *Community Mental Health Journal*, 46(5), 494-499.
- Carpenter, K. M., Martinez, D., Vadhan, N. P., Barnes-Holmes, D., & Nunes, E. V. (2013). Measures of attentional bias and relational responding are associated with behavioral treatment outcome for cocaine dependence. *American Journal of Drug and Alcohol Abuse*, 38(2), 146-154.
- Chadwick, P., & Birchwood, M. (1994). The omnipotence of voices. A cognitive approach to auditory hallucinations. *British Journal of Psychiatry*, 164(2), 190-201.
- Chadwick, P., & Birchwood, M. (1995). The omnipotence of voices II: The beliefs about voices questionnaire (BAVQ). *The British Journal of Psychiatry*, 166, 773-776.

- Chadwick, P., Lees, S., & Birchwood, M. (2000). The revised Beliefs About Voices Questionnaire (BAVQ-R). *The British Journal of Psychiatry*, *177*, 229–232.
- Choong, C., Hunter, M. D., & Woodruff, P. W. (2007). Auditory hallucinations in those populations that do not suffer from schizophrenia. *Current Psychiatry Reports*, *9*(3), 206-212.
- Corrigan, P. (2004). How stigma interferes with mental health care. *American Psychologist*, *59*(7), 614-625.
- Corstens, D., Longden, E., McCarthy-Jones, S., Waddingham, R., & Thomas, N. (2014). Emerging perspectives from the hearing voices movement: implications for research and practice. *Schizophrenia Bulletin*, *40*(4), 285–294.
- Crisp, A. H., Gelder, M. G., Rix, S., Mertzner, I., & Rowlands, O. J. (2000). Stigmatisation of people with mental illnesses. *British Journal of Psychiatry*, *177*, 4-7.
- Daalman, K., Boks, M. P., Diederens, K. M., de Weijer, A. D., Blom, J. D., Kahn, R. S., & Sommer, I. E. (2011). The same or different? A phenomenological comparison of auditory verbal hallucinations in healthy and psychotic individuals. *Journal of Clinical Psychiatry*, *72*(3), 320-325.
- Dearing, K. S., & Steadman, S. (2009). Enhancing intellectual empathy: The lived experience of voice simulation. *Perspectives in Psychiatric Care*, *45*(3), 173–182.
- De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology*, *50*, 77-85.

- De Houwer, J. (2006). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.), *The Handbook of Implicit Cognition and Addiction* (pp. 11-28). CA: Sage Publishers.
- Deegan, P. (1996). *Hearing Voices that are Distressing: A Training and Simulated Experience*. National Empowerment Center.
- Diederer, K. M. J., Van Lutterveld, R., & Sommer, I. E. C. (2012). Neuroimaging of voice hearing in non-psychotic individuals: a mini review. *Frontiers in Human Neuroscience*, 6, 1–5.
- Dorahy, M. J., Shannon, C., Seagar, L., Corr, M., Stewart, K., Hanna, D., ... Middleton, W. (2009). Auditory hallucinations in dissociative identity disorder and schizophrenia with and without a childhood trauma history: Similarities and differences. *Journal of Nervous and Mental Disease*, 3(12), 892-898.
- Drake, R., Haddock, G., Tarrier, N., Bentall, R., & Lewis, S. (2007). The Psychotic Symptom Rating Scales (PSYRATS): Their usefulness and properties in first episode psychosis. *Schizophrenia research*, 89(1-3), 119–22.
- Eaton, W. W., Romanoski, A., Anthony, J. C., & Nestadt, G. (1991). Screening for psychosis in the general population with a explicit interview. *The Journal of Nervous and Mental Disease*, 179, 689-693.
- Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology*, 83(6), 1441-1455.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, 69, 1013-1027.

- Fischer, S. (2013). Assessing the test-retest reliability of the IRAP in the context of sexuality. (Unpublished bachelor thesis). National University of Ireland Maynooth, Ireland.
- Foody, M., Barnes-Holmes, Y., & Barnes-Holmes, D. (2012). The role of self in Acceptance and Commitment Therapy (ACT). In: L. McHugh, & I. Stewart (Eds.), *The self and perspective taking: Research and applications*. Oakland, CA: New Harbinger.
- Gauntlett-Gilbert, J., & Kuipers, E. (2005). Visual hallucinations in psychiatric conditions: appraisals and their relationship to distress. *British Journal of Clinical Psychology, 44*, 77–87.
- Goldstone, E., Farhall, J., & Ong, B. (2011). Life hassles, experiential avoidance and distressing delusional experiences. *Behaviour Research and Therapy 49*, 260–266.
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3–25.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality & Social Psychology, 85*(2), 197-216.
- Greenwald, A. G., Poehlman, A., Uhlmann, E., & Banaji, M. R. (2009). Understanding and interpreting the Implicit Association Test III: Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*(1), 17-41.

- Haddock, G., McCarron, J., Tarrier, N., & Farragher, E. B. (1999). Scales to measure dimensions of hallucinations and delusions: The Psychotic Symptom Rating Scales (PSYRATS). *Psychological Medicine*, *29*, 879–889.
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational Frame Theory: A post Skinnerian account of human language and cognition*. New York: Plenum Press.
- Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, *45*, 227–239.
- Hill, K., Varese, F., Jackson, M., & Linden, D. E. (2012). The relationship between metacognitive beliefs, auditory hallucinations, and hallucination-related distress in clinical and non-clinical voice-hearers. *British Journal of Clinical Psychology*, *51*, 434–447.
- Honig, A., Romme, M. A. J., Ensink, B. J., Escher, S. D., Pennings, M. H. A., & Devries, M. W. (1998). Auditory hallucinations: A comparison between patients and non-patients. *The Journal of Nervous and Mental Disease*, *189*(10), 646-651.
- Hollon, S., & Kendall, P. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research*, *4*(4), 383–395.
- Hussey, I., & Barnes-Holmes, D. (2012). The Implicit Relational Assessment Procedure as a measure of implicit depression and the role of psychological flexibility. *Cognitive and Behavioral Practice*, *19*(4), 573-583.

- Hussey, I., McEnteggart, C., Nicholson, E., Kavanagh, D., Thompson, M., Barnes-Holmes, D., & Barnes-Holmes, Y. (under submission). Empirical ethics: Initial psychometric properties of the absolute and minimal risk scales. *British Journal of Psychology*.
- Jackson, L. J., Hayward, M., & Cooke, A. (2011). Developing positive relationships with voices: A preliminary Grounded Theory. *International Journal of Social Psychiatry, 57*, 487–495.
- Jenner, J. A., Rutten, S., Beuckens, J., Boonstra, N., & Sytema, S. (2008). Positive and useful auditory vocal hallucinations: prevalence, characteristics, attributions, and implications for treatment. *Acta psychiatrica Scandinavica, 118*(3), 238–45.
- Johns, L. C., Kompus, K., Connell, M., Humpston, C., Lincoln, T. M., Longden, E., ... Laroi, F. (2014). Auditory verbal hallucinations in persons with and without a need for care. *Schizophrenia Bulletin, 40*(4), S255-S264.
- Johns, L. C., & Van Os, J. (2001). The continuity of psychotic experiences in the general population. *Clinical Psychology Review, 21*(8), 1125–1141.
- Kalyanaraman, S., Penn, D., Ivory, J., & Judge, A. (2010). The virtual doppelganger: Effects of a virtual reality simulator on perceptions of schizophrenia. *Journal of Nervous and Mental Disorders, 198*(6), 437-443.
- Kidd, L. I., Tusaie, K. R., Morgan, K. I., Preebe, L., Garrett, M. (2015). Mindful teaching practice: Lessons learned through a hearing voices simulation. *Issues in Mental Health Nursing, 36*(2), 112-117.
- Kirk, S., & Kutchins, H. (1994). The myth of the reliability of the DSM. *Journal of Mind and Behaviour, 15*, 71–86.

- Kim, S. H., Jung, H. Y., Hwang, S. S., Chang, J. S., Kim, Y., Ahn, Y. M., & Kim, Y. S. (2010). The usefulness of an explicit questionnaire measuring auditory verbal hallucinations. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *34*(6), 968-973.
- Kleinman A., Eisenberg L., & Good B. (1978). Culture, illness, and care: clinical lessons from anthropological and cross-cultural research. *Annual Guide to International Medicine*, *88*, 251–88.
- Konings, M., Bak, M., Hanssen, M., Van Os, J., & Krabbendam, L. (2006). Validity and reliability of the CAPE: An explicit instrument for the measurement of psychotic experiences in the general population. *Acta psychiatrica Scandinavica*, *114*(1), 55–61.
- Krabbendam, L., Myin-Germeys, I., Bak, M., & Van Os, J. (2005). Explaining transitions over the hypothesized psychosis continuum. *Australian and New Zealand Journal of Psychiatry*, *39*, 180–186.
- Lauber, C., Nordt, C., Falcato, L., & Rössler, W. (2003). Do people recognise mental illness? Factors influencing mental health literacy. *European Archives of Psychiatry and Clinical Neuroscience*, *253*, 248-251.
- Levenson, H. (1973). Multidimensional locus of control in psychiatric patients. *Journal of Consulting and Clinical Psychology*, *41*, 397-404.
- Lincoln, T. M., Arens, E., Berger, C., & Rief, W. (2008). Can antistigma campaigns be improved? A test of the impact of biogenetic vs. psychosocial causal explanations on implicit and explicit attitudes to schizophrenia. *Schizophrenia Bulletin*, *34*(5), 984- 994.

- Link, B. G., Cullen, F. T., Struening, E., Shrout, P. E., & Dohrenwend, B. P. (1989). A modified labelling theory approach to mental disorders: An empirical assessment. *American Sociological Review*, *54*, 400-423.
- Link, B. G., & Phelan, J. C. (2001). Conceptualizing stigma. *Annual Review of Sociology*, *27*, 104-106.
- Link, B. G., Yang, L. H., Phelan, J. C., & Collins, P. Y. (2004). Measuring mental illness stigma. *Schizophrenia Bulletin*, *30*(3), 511-541.
- Longden, E., Corstens, D., Escher, S., & Romme, M. (2012). Voice hearing in a biographical context: a model for formulating the relationship between voices and life history. *Psychosis*, *4*, 224–234
- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales* (2nd ed.). Sydney: The Psychology Foundation of Australia.
- Luciano, C., Valdivia-Salas, S., Ruiz, F. J., Rodriguez-Valverde, M., Barnes-Holmes, D., Dougher, M. J. ... Gutierrez, O. (2013). Extinction of aversive eliciting functions as an analogue of exposure to conditioned fear: Does it alter avoidance responding? *Journal of Contextual Behavioral Science*, *2*(3-4), 120–134.
- Mak, W. W. S., Poon, C. Y. M., Pun, L. Y. K., & Cheung, S. F. (2007). Meta-analysis of stigma and mental health. *Social Science & Medicine*, *65*(2), 245-261.
- Masuda, A., Price, M., Anderson, P. L., Schmertz, S. K., & Calamaras, M. R. (2009). The role of psychological flexibility in mental health stigma and psychological distress for the stigmatizer. *Journal of Social and Clinical Psychology*, *28*(10), 1244-1262.

- Mawson, A., Berry, K., Murray, C., & Hayward, M. (2011). Voice hearing within the context of hearers' social worlds: An interpretative phenomenological analysis. *Psychology and Psychotherapy: Theory, Research and Practice*, 84, 256–272.
- Mawson, A., Cohen, K., & Berry, K. (2010). Reviewing evidence for the cognitive model of auditory hallucinations: The relationship between cognitive voice appraisals and distress during psychosis. *Clinical Psychology Review*, 30, 248–258.
- McCulloch, Y., Clare, L., Howard, R., & Peters, E. (2006). Psychological processes underlying delusional thinking in late-onset psychosis: a preliminary investigation. *International Journal Of Geriatric Psychiatry*, 21(8), 768-777.
- Mersh, L., Jones, F., & Oliver, J. (in press). Mindfulness, self-stigma and social functioning in first episode psychosis: A brief report. *Psychosis*.
- Morris, E. M., Garety, P., & Peters, E. (2014). Psychological flexibility and nonjudgemental acceptance in voice hearers: relationships with omnipotence and distress. *The Australian and New Zealand journal of psychiatry*, 48(12), 1150-1162.
- Morrison, A. P. (2001). The interpretation of intrusions in psychosis: An integrative cognitive approach to hallucinations and delusions. *Behavioural and Cognitive Psychotherapy*, 29, 257-276.
- Morrison, M., D'man, A. F., & Drumheller, A. (1994). Multidimensional locus of control and attitudes toward mental illness. *Perceptual and Motor Skills*, 78(3), 1281-1282.
- Morrison, A. P., Nothard, S., Bowe, S. E., & Wells, A. (2004). Interpretations of voices in patients with hallucinations and non-patient controls: A comparison

- and predictors of distress in patients. *Behaviour Research and Therapy*, 42, 1315-1323.
- Mossaheb, N., Becker, J., Schaefer, M., Klier, C., Schloegelhofer, M., Papageorgiou, K., & Amminger, G. (2012). The Community Assessment of Psychic Experience (CAPE) questionnaire as a screening-instrument in the detection of individuals at ultra-high risk for psychosis. *Schizophrenia research*, 141(2-3), 210–4.
- Nicholson, E., & Barnes-Holmes, D. (2012a). The Implicit Relational Assessment Procedure (IRAP) as a measure of spider fear. *The Psychological Record*, 62, 263-277.
- Nicholson, E., & Barnes-Holmes, D. (2012b). Developing an implicit measure of disgust propensity and disgust sensitivity: Examining the role of implicit disgust propensity and sensitivity in obsessive-compulsive tendencies. *Journal of Behavioural Therapy and Experimental Psychiatry*, 43(3), 922-930.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231-259.
- Nock, M. K., Park, J. L., Finn, C. T., Deliberto, T. L., Dour, H. J., & Banaji, M. R. (2010). Measuring the “suicidal mind”: Implicit cognition predicts suicidal behavior. *Psychological Science*, 21, 511–517.
- O’Brien, P., & Johns, L. (2013). A graded exposure intervention for distressing visual hallucinations in schizophrenia. *Behavioural and Cognitive Psychotherapy*, 41, 495-499.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in*

psychological and educational measurement (pp. 49–69). Mahwah NJ: Erlbaum.

Peris, T. S., Teachman, B. A., & Nosek, B. A. (2008). Implicit and explicit stigma of mental illness: Links to clinical care. *Journal of Nervous and Mental Disease, 196*(10), 752-760.

Peters, E. R., Joseph, S. A., & Garety, P. A. (1999). Measurement of delusional ideation in the normal population: Introducing the PDI (Peters et al. Delusions Inventory). *Schizophrenia Bulletin, 25*(3), 553-576.

Peters, E. R., Williams, S. L., Cooke, M. A., & Kuipers, E. (2012). It's not what you hear, it's the way you think about it: Appraisals as determinants of affect and behaviour in voice hearers. *Psychological Medicine, 42*(7), 1507-1514.

Psychological Society of Ireland (PSI). (2011). *Code of Professional Ethics of the Psychological Society of Ireland*. Accessed from: <http://www.psychologicalsociety.ie/find-apsychologist/PSI%202011-12%20Code%20of%20Ethics.pdf>.

Ratcliff, K., Farhall, J., & Shawyer, F. (2010). Auditory Hallucinations : A Review of Assessment Tools. *Clinical Psychology and Psychotherapy, 18*, 524-534.

Reavley, N. J., & Jorn, A. F. (2011). Stigmatizing attitudes towards people with mental disorders: findings from an Australian national survey of mental health literacy and stigma. *Australian & New Zealand Journal of Psychiatry, 45*, 1086-1093.

Romme, M. A., & Escher, A. D. (1989). Hearing voices. *Schizophrenia Bulletin, 15*(2), 209- 216.

Rosenfarb, I. S. (2013). A Functional Analysis of Schizophrenia. *The Psychological Record, 63*, 929–946.

- Rössler, W., Riecher-Rössler, A., Angst, J., Murray, R., Gamma, A., Eich, D., ...
Gross, V. A. (2007). Psychotic experiences in the general population: A
twenty-year prospective community study. *Schizophrenia Research*, 92(1), 1-
14.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of
reinforcement. *Psychological Monographs*, 33(1), 300-303.
- Rüsch, N., Corrigan, P. W., Todd, A. R., & Bodenhausen, G. V. (2010). Implicit
self-stigma in people with mental illness. *Journal of Nervous and Mental
Disorders*, 198(2), 150-153.
- Ryu, S. (2001). 정신장애인에 영향을 주는 제 변인 (The effect of variables on the
attitude toward mental illness). 고려대학교 석사학위 청구논문.
- Sartorius, N., Jablensky, A., Korten, A., Ernberg, G., Anker, M., Cooper, J. E., &
Day, R. (1986). Early manifestations and first contact incidence of
schizophrenia in different cultures. *Psychological Medicine*, 16(4), 909-928.
- Schulze, B. (2007). Stigma and mental health professionals: A review of the
evidence on an intricate relationship. *International Review of Psychiatry*,
19(2), 137-155.
- Shawyer, F., Ratcliff, K., Mackinnon, A., Farhall, J., Hayes, S., & Copolov, D.
(2007). The Voices Acceptance and Action Scale (VAAS): Pilot Data.
Journal of Clinical Psychology, 63(6), 593-606.
- Sideras, S., McKenzie, G., Noone, J., Dieckmann, N., & Allen, T. L. (2015). Impact
of a simulation on nursing students' attitudes toward schizophrenia. *Clinical
Simulation in Nursing*, 11(2), 134-141.

- Slotema, C. W., Daalman, K., Blom, J. D., Diederer, K. M., Hoek, H. W., & Sommer, I. E. C. (2012). Auditory verbal hallucinations in patients with borderline personality disorder are similar to those in schizophrenia. *Psychological Medicine, 42*(9), 1873-1878.
- Sommer, I. E., Daalman, K., Rietkerk, T., Diederer, K. M., Bakker, S., Wijkstra, J., & Boks, M. P. (2010). Healthy individuals with auditory verbal hallucinations: Who are they? 27 psychiatric assessments of a selected sample of 103 subjects. *Schizophrenia Bulletin, 36*(3), 633-641.
- Soler, M. J., Ruiz, J. C., Vargas, M., Dasí, C., & Fuentes, I. (2011). Perceptual priming in schizophrenia evaluated by word fragment and word stem completion, *Psychiatry Research, 30*(2-3), 167-171.
- Stacy, A. W. (1997). Memory activation and expectancy as prospective predictors of alcohol and marijuana use. *Journal of Abnormal Psychology, 106*, 61-73.
- Stefanis, N., Hanssen, M., Smirnis, N., Avramopoulos, D., Evdokimidis, I., Stefanis, C., ... Van Os, J. (2002). Evidence that three dimensions of psychosis have a distribution in the general population. *Psychological medicine, 32*(2), 347-58.
- Stier, A., & Hinshaw, S. P. (2007). Explicit and implicit stigma against individuals with mental illness. *Australian Psychologist, 42*(2), 106-117.
- Szasz, T. (1960). The Myth of Mental Illness. *American Psychologist, 15*(2), 113-118.
- Taylor, S. M., Dear, M. J., & Hall, G. B. (1979). Attitudes towards the mentally ill and reactions to mental health facilities. *Social Sciences and Medicine, 13*, 281-290.

- Taylor, S. M., & Dear, M. J. (1981). Scaling community attitudes toward the mentally ill. *Schizophrenia Bulletin*, 7 (2), 225-240.
- Teachman, B. A., & Woody, S. (2004). Staying tuned to research in implicit cognition: Relevance for clinical practice with anxiety disorders. *Cognitive and Behavioral Practice*, 11, 149-159.
- Teachman, B. A., Wilson, H. G., & Komarovskaya, I. (2006). Implicit and explicit stigma of mental illness in diagnosed and healthy samples. *Journal of Social & Clinical Psychology*, 25(1), 75-95.
- Tien, A. Y. (1991). Distribution of hallucinations in the population. *Social Psychiatry and Psychiatric Epidemiology*, 26, 287-292.
- Udachina, A., Varese, F., Myin-Germeys, I., & Bentall, R. (2012). The role of experiential avoidance in paranoid delusions: An experience sampling study. *The British journal of Clinical Psychology*, 53(4), 422-432.
- Vahey, N.A, Nicholson, E., & Barnes-Holmes, D. (2015). A meta-analysis of criterion effects for the Implicit Relational Assessment Procedure (IRAP) in the clinical domain. *Journal of Behavior Therapy and Experimental Psychiatry*, 48, 59-65.
- Vázquez, C., Díez-Alegría, C., Hernández-Lloreda, M. J., & Moreno, M. N. (2008). Implicit and explicit self-schema in active deluded, remitted deluded, and depressed patients, *Journal of Behavior Therapy and Experimental Psychiatry*, 39, 587-599.
- Van Os, J., Gilvarry, C., Bale, R., Van Horn, E., Tattan, T., White, I., & Murray, R. (1999). A comparison of the utility of dimensional and categorical representations of psychosis. *Psychological Medicine*, 29(3), 595–606.

- Vaughan, S., & Fowler, D. (2004). The distress experienced by voice hearers is associated with the perceived relationship between the voice hearer and the voice. *British Journal of Clinical Psychology, 43*, 143–153.
- van der Gaag, M., Hageman, M. C., & Birchwood, M. (2003). Evidence for a cognitive model of auditory hallucinations. *Journal of Nervous and Mental Disease, 191*, 542–545.
- Ward, T. A., Gaynor, K. J., Hunter, M. D., Woodruff, P. W. R., Garety, P. A., & Peters, E. R. (2014). Appraisals and responses to experimental symptom analogues in clinical and nonclinical individuals with psychotic experiences. *Schizophrenia Bulletin, 40*(4), 845-855.
- White, R., Gumley, A., McTaggart, J., Rattrie, L., McConville, D., Cleare, S., & Mitchell, G. (2011). A feasibility study of Acceptance and Commitment Therapy for emotional dysfunction following psychosis. *Behaviour research and therapy, 49*(12), 901-907.
- Wieland, D., Levine, C., & Smith, J. (2015). Hearing distressing voices clinical simulation: “Life changing” experiences of psychiatric-mental health nursing students. *Journal of Psychosocial Nursing and Mental Health Services, 52*(10), 42-51.
- Wiffen, B. D. R., O'Connor, J. A., Russo, M., Falcone, M. A., Joseph, C., Kolliakou, A., ... David, A.S. (2014). Do psychosis patients with poor insight show implicit awareness on the emotional Stroop task? *Psychopathology, 47*(2), 93-100.
- Wilson, J., Azzopardi, W., Sager, S., Gould, B., Conroy, S., Deegan, P., & Archie, S. (2009). A narrative study of the experiences of nursing students who have

participated in the “Hearing Voices that are Distressing” simulation.

International Journal of Nursing Education Scholarship, 6(1), 1–15.

Wilson, T. D., & Dunn, E. W. (2004). Self-knowledge: Its limits, value, and potential for improvement. *Annual Review of Psychology*, 55, 493-518.

Woods, A., Jones, N., Alderson-Day, B., Callard, F., & Fernyhough, C. (2015). Experiences of hearing voices: Analysis of a novel phenomenological survey.

The Lancet Psychiatry, 2(4), pp. 323-331.

Appendices

Appendix 1: The Community Attitudes to Mentally Ill Questionnaire

CAMI

The following statements express various opinions about mental illness. Please **circle** the response that most accurately describes your FIRST reaction to each statement.

1. As soon as a person shows signs of mental disturbance, they should be hospitalized.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

2. More money should be spent on the care and treatment of adults with mental illness.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

3. An adult with mental illness should be isolated from the rest of the community.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

4. The best therapy for many adults with mental illness is to be part of a normal community.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

5. Mental illness is an illness like any other.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

6. Adults with mental illness are a burden on society.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

7. Adults with mental illness are far less of a danger than most people think.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

8. Locating mental health facilities in a residential area downgrades the neighborhood.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

9. There is something about adults with mental illness that makes it easy to tell them from normal people.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

10. Adults with mental illness have for too long been the subject of ridicule.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

11. A person would be foolish to marry someone who has suffered from mental illness, even though they seem fully recovered.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

12. As far as possible mental health services should be provided through community-based facilities.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

13. Less emphasis should be placed on protecting the public from adults with mental illness.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

14. Increased spending on mental health services is a waste of tax money.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

15. No one has the right to exclude adults with mental illness from their neighborhood.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

16. Having adults with mental illness living within residential neighborhoods might be good therapy, but the risks to residents are too great.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

17. Adults with mental illness need the same kind of control and discipline as a young child.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

18. We need to adopt a far more tolerant attitude toward adults with mental illness in our society.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

19. I would not want to live next door to someone who has been mentally ill.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

20. Residents should accept the location of mental health facilities in their neighborhood to serve the needs of the local community.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

21. Adults with mental illness should not be treated as outcasts of society.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

22. There are sufficient existing services for adults with mental illness.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

23. Adults with mental illness should be encouraged to assume the responsibilities of normal life.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

24. Local residents have good reason to resist the location of mental health services in their neighborhood.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

25. The best way to handle adults with mental illness is to keep them behind locked doors.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

26. Our mental hospitals seem more like prisons than like places where adults with mental illness can be cared for.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

27. Anyone with a history of mental illness should be excluded from working in government.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

28. Locating mental health services in residential neighborhoods does not endanger local residents.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

29. Mental hospitals are an outdated means of treating adults with mental illness.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

30. Adults with mental illness do not deserve our sympathy.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

31. Adults with mental illness should not be denied their individual rights.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

32. Mental health facilities should be kept out of residential neighborhoods.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

33. One of the main causes of mental illness is a lack of self-discipline and will power.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

34. We have the responsibility to provide the best possible care for adults with mental illness.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

35. Adults with mental illness should not be given any responsibility.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

36. Residents have nothing to fear from people coming into their neighborhood to obtain mental health services.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

37. Virtually anyone can become mentally ill.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

38. It is best to avoid anyone who has mental problems.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

39. Most people who were once patients in a mental hospital can be trusted as baby sitters.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

40. It is frightening to think of people with mental problems living in residential neighborhoods.

Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

Appendix 3: Depression Anxiety and Stress Scales

DASS 21

Please read each statement and circle a number 0, 1, 2 or 3 which indicates how much the statement applied to you **over the past week**

	0	1	2	3
	Did not apply to me at all	Applied to me to some degree, or some of the time	Applied to me to a considerable degree, or a good part of time	Applied to me very much, or most of the time
1	I found it hard to wind down			0 1 2 3
2	I was aware of dryness of my mouth			0 1 2 3
3	I couldn't seem to experience any positive feeling at all			0 1 2 3
4	I experienced breathing difficulty (eg, excessively rapid breathing, breathlessness in the absence of physical exertion)			0 1 2 3
5	I found it difficult to work up the initiative to do things			0 1 2 3
6	I tended to over-react to situations			0 1 2 3
7	I experienced trembling (eg, in the hands)			0 1 2 3
8	I felt that I was using a lot of nervous energy			0 1 2 3
9	I was worried about situations in which I might panic and make a fool of myself			0 1 2 3
10	I felt that I had nothing to look forward to			0 1 2 3
11	I found myself getting agitated			0 1 2 3
12	I found it difficult to relax			0 1 2 3
13	I felt down-hearted and blue			0 1 2 3
14	I was intolerant of anything that kept me from getting on with what I was doing			0 1 2 3
15	I felt I was close to panic			0 1 2 3
16	I was unable to become enthusiastic about anything			0 1 2 3
17	I felt I wasn't worth much as a person			0 1 2 3
18	I felt that I was rather touchy			0 1 2 3
19	I was aware of the action of my heart in the absence of physical exertion (eg, sense of heart rate increase, heart missing a beat)			0 1 2 3
20	I felt scared without any good reason			0 1 2 3
21	I felt that life was meaningless			0 1 2 3

Appendix 4: Consent Form

Consent Form

Participant

I consent to participate in an experimental psychology study being run by Ciara McEnteggart. I understand and consent to the following:

- The experiment will last a maximum time of one hour.
- I understand that the experiment has two parts including questionnaires and a computer task. I understand that I may not have to complete both parts of the experiment.
- I am free to terminate my participation in the study at any time and may withdraw the data obtained from my participation, if I so wish.
- I understand that the experiment does not offer any therapeutic intervention.
- I understand that I participate under my own volition and that my participation will not have any effect on my subsequent academic results. I also understand that no monetary remuneration will result from participation.
- I understand that the data collected will be safeguarded in a code protected computer system, and any raw data will be locked in a cabinet, for a period of five years, after which it will be destroyed. Up until this period, I may have access to the data collected.
- I understand that the data will be combined, analysed, and may be presented at International Conferences, or submitted to international journals for publication.

I have received this information in an understandable way. All my questions have been answered.

Please print and sign your name below if you are willing to abide fully by the conditions stated above.

Name: _____ **(Please print in block capitals)**

Signature: _____

Date: _____

Experimenter

I, Ciara McEnteggart, as primary experimenter, I can confirm that all the necessary safety precautions have been taken.

Signature of experimenter: _____

Date: _____

Contact Details

Researcher: Ciara McEnteggart, Department of Psychology, 2nd Floor, John Hume Building, NUI Maynooth, Maynooth, Co. Kildare
Tel: +353 87 975 8633
Email: ciara.mcenteggart@nuim.ie

Supervisor: Dr. Yvonne Barnes-Holmes, Department of Psychology, 2nd Floor, John Hume Building, NUI Maynooth, Maynooth, Co. Kildare.
Tel: +3531 708 6080
Email: yvonne.barnes-holmes@nuim.ie

If during your participation in this study you feel the information and guidelines that you were given have been neglected or disregarded in any way, or if you are unhappy about the process, please contact the Secretary of the National University of Ireland Maynooth Ethics Committee at research.ethics@nuim.ie or +353 (0)1 708 6019. Please be assured that your concerns will be dealt with in a sensitive manner.

Appendix 5: Illness Questionnaire

Illness Questionnaire

Please rate the following statements by either circling Yes or No.

- | | | |
|--|-----|----|
| 1. Do you think depression is an illness? | Yes | No |
| 2. Do you think anxiety is an illness? | Yes | No |
| 3. Do you think phobias are an illness? | Yes | No |
| 4. Do you think worry is an illness? | Yes | No |
| 5. Do you think obsessions are an illness? | Yes | No |
| 6. Do you think panic is an illness? | Yes | No |

Appendix 6: Auditory Hallucinations Rating Scale

Auditory hallucinations Rating Scale

1 Frequency

- 0 Voices not present or present less than once a week
- 1 Voices occur for at least once a week
- 2 Voices occur at least once a day
- 3 Voices occur at least once a hour
- 4 Voices occur continuously or almost continuously i.e. stop for only a few seconds or minutes

2 Duration

- 0 Voices not present
- 1 Voices last for a few seconds, fleeting voices
- 2 Voices last for several minutes
- 3 Voices last for at least one hour
- 4 Voices last for hours at a time

3 Location

- 0 No voices present
- 1 Voices sound like they are inside head only
- 2 Voices outside the head, but close to ears or head. Voices inside the head may also be present
- 3 Voices sound like they are inside or close to ears and outside head away from ears
- 4 Voices sound like they are from outside the head only

4 Loudness

- 0 Voices not present
- 1 Quieter than own voice, whispers.
- 2 About same loudness as own voice
- 3 Louder than own voice
- 4 Extremely loud, shouting

5 Beliefs re-origin of voices

- 0 Voices not present
- 1 Believes voices to be solely internally generated and related to self
- 2 Holds < 50% conviction that voices originate from external causes
- 3 Holds ~ 50% conviction (but < 100%) that voices originate from external causes
- 4 Believes voices are solely due to external causes (100% conviction)

6 Amount of negative content of voices

- 0 No unpleasant content
- 1 Occasional unpleasant content (< 10%)
- 2 Minority of voice content is unpleasant or negative (< 50%)
- 3 Majority of voice content is unpleasant or negative (> 50%)
- 4 All of voice content is unpleasant or negative

7 Degree of negative content

- 0 Not unpleasant or negative
- 1 Some degree of negative content, but not personal comments relating to self or family e.g. swear words or comments not directed to self, e.g. 'the milkman's ugly'
- 2 Personal verbal abuse, comments on behavior e.g. ' shouldn't do that or say that ,
- 3 Personal verbal abuse relating to self-concept e.g. 'you're lazy, ugly, mad, perverted ,
- 4 Personal threats to self e.g. threats to harm self or family, extreme instructions or commands to harm self or others

8 Amount of distress

- 0 Voices not distressing at all
- 1 Voices occasionally distressing, majority not distressing (< 10%)
- 2 Minority of voices distressing (< 50%)
- 3 Majority of voices distressing, minority not distressing (~ 50%)
- 4 Voices always distressing

9 Intensity of distress

- 0 Voices not distressing at all
- 1 Voices slightly distressing
- 2 Voices are distressing to a moderate degree
- 3 Voices are very distressing, although subject could feel worse
- 4 Voices are extremely distressing, feel the worst he/she could possibly feel

10 Disruption to life caused by voices

- 0 No disruption to life, able to maintain social and family relationships (if present)
- 1 Voices causes minimal amount of disruption to life e.g. interferes with concentration although able to maintain daytime activity and social and family relationships and be able to maintain independent living without support
- 2 Voices cause moderate amount of disruption to life causing some disturbance to daytime activity and/or family or social activities. The patient is not in hospital although may live in supported accommodation or receive additional help with daily living skills
- 3 Voices cause severe disruption to life so that hospitalisation is usually necessary . The patient is able to maintain some daily activities, self-care and relationships while in hospital. The patient may also be in supported accommodation but experiencing severe disruption of life in terms of activities, daily living skills and/or relationships
- 4 Voices cause complete disruption of daily life requiring hospitalization. The patient is unable to maintain any daily activities and social relationships. Self-care is also severely disrupted.

11 Controllability of voices

- 0 Subject believes they can have control over the voices and can always bring on or dismiss them at will
- 1 Subject believes they can have some control over the voices on the majority of occasions
- 2 Subject believes they can have some control over their voices approximately half of the time
- 3 Subject believes they can have some control over their voices but only occasionally. The majority of the time the subject experiences voices which are uncontrollable
- 4 Subject has no control over when the voices occur and cannot dismiss or bring them on at all

Appendix 7: Community Assessment of Psychic Experiences Questionnaire

1. Do you ever feel sad?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 2

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

2. Do you ever feel as if people seem to drop hints about you or say things with a double meaning?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 3

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

3. Do you ever feel that you are not a very animated person?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 4

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

4. Do you ever feel that you are not much of a talker when you are conversing with other people?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 5

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

5. Do you ever feel as if things in magazines or on TV were written especially for you?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 6

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

6. Do you ever feel as if some people are not what they seem to be?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 7

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

7. Do you ever feel as if you are being persecuted in some way?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 8

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

8. Do you ever feel that you experience few or no emotions at important events?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 9

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

9. Do you ever feel pessimistic about everything?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 10

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

10. Do you ever feel as if there is a conspiracy against you?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 11

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

11. Do you ever feel as if you are destined to be someone very important?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 12

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

12. Do you ever feel as if there is no future for you?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 13

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

13. Do you ever feel that you are a very special or unusual person?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 14

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

14. Do you ever feel as if you do not want to live anymore?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 15

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

15. Do you ever think that people can communicate telepathically?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 16

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

16. Do you ever feel that you have no interest to be with other people?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 17

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

17. Do you ever feel as if electrical devices such as computers can influence the way you think?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 18

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

18. Do you ever feel that you are lacking in motivation to do things?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 19

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

19. Do you ever cry about nothing?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 20

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed
20. Do you believe in the power of witchcraft, voodoo or the occult?
Never Sometimes Often Nearly always

If you ticked "never", please go to question 21

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

21. Do you ever feel that you are lacking in energy?
Never Sometimes Often Nearly always

If you ticked "never", please go to question 22

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

22. Do you ever feel that people look at you oddly because of your appearance?
Never Sometimes Often Nearly always

If you ticked "never", please go to question 23

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

23. Do you ever feel that your mind is empty?
Never Sometimes Often Nearly always

If you ticked "never", please go to question 24

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

24. Do you ever feel as if the thoughts in your head are being taken away from you?
Never Sometimes Often Nearly always

If you ticked "never", please go to question 25

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

25. Do you ever feel that you are spending all your days doing nothing?
Never Sometimes Often Nearly always

If you ticked "never", please go to question 26

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

26. Do you ever feel as if the thoughts in your head are not your own?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 27

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

27. Do you ever feel that your feelings are lacking in intensity?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 28

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

28. Have your thoughts ever been so vivid that you were worried other people would hear them?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 29

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

29. Do you ever feel that you are lacking in spontaneity?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 30

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

30. Do you ever hear your own thoughts being echoed back to you?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 31

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

31. Do you ever feel as if you are under the control of some force or power other than yourself?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 32

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

32. Do you ever feel that your emotions are blunted?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 33

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

33. Do you ever hear voices when you are alone?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 34

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

34. Do you ever hear voices talking to each other when you are alone?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 35

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

35. Do you ever feel that you are neglecting your appearance or personal hygiene?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 36

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

36. Do you ever feel that you can never get things done?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 37

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

37. Do you ever feel that you have only few hobbies or interests?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 38

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

38. Do you ever feel guilty?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 39

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

39. Do you ever feel like a failure?

Never Sometimes Often Nearly always

If you ticked "never", please go to question 40

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

40. Do you ever feel tense?

Never Sometimes Often Nearly always

If you ticked "never", you are now ready

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

41. Do you ever feel as if a double has taken the place of a family member, friend or acquaintance?

Never Sometimes Often Nearly always

If you ticked "never", you are now ready

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

42. Do you ever see objects, people or animals that other people cannot see?

Never Sometimes Often Nearly always

If you ticked "never", you are now ready

If you ticked "sometimes" , "often" or "nearly always" please indicate how distressed you are by this experience:

Not distressed A bit distressed Quite distressed Very distressed

Appendix 8: Acceptance and Action Questionnaire II

AAQ-II

Below you will find a list of statements. Please rate how true each statement is for you by circling a number next to it. Use the scale below to make your choice.

1	2	3	4	5	6	7
never true	very seldom true	seldom true	sometimes true	frequently true	almost always true	always true

1. My painful experiences and memories make it difficult for me to live a life that I would value.	1	2	3	4	5	6	7
2. I'm afraid of my feelings.	1	2	3	4	5	6	7
3. I worry about not being able to control my worries and feelings.	1	2	3	4	5	6	7
4. My painful memories prevent me from having a fulfilling life.	1	2	3	4	5	6	7
5. Emotions cause problems in my life.	1	2	3	4	5	6	7
6. It seems like most people are handling their lives better than I am.	1	2	3	4	5	6	7
7. Worries get in the way of my success.	1	2	3	4	5	6	7

Appendix 9: Automatic Thoughts Questionnaire

Instructions: Listed below are a variety of thoughts that pop into people's heads. Please read each thought and indicate how frequently, if at all, the thought occurred to you *over the last week*.

Please circle a response on the LEFT side of the sheet using the **FREQUENCY** scale:

1 = not at all 2 = sometimes 3 = moderately often 4 = often 5 = all the time

Then, please indicate how strongly, if at all, you tend to believe that thought, when it occurs. Please circle a response on the RIGHT side of the sheet using the **DEGREE OF BELIEF** scale:

1 = not at all 2 = somewhat 3 = moderately 4 = very much 5 = totally

Frequency	Item	Degree of Belief
1 2 3 4 5	1.) I feel like I'm up against the world.	1 2 3 4 5
1 2 3 4 5	2.) I'm no good.	1 2 3 4 5
1 2 3 4 5	3.) Why can't I ever succeed?	1 2 3 4 5
1 2 3 4 5	4.) No one understands me.	1 2 3 4 5
1 2 3 4 5	5.) I've let people down.	1 2 3 4 5
1 2 3 4 5	6.) I don't think I can go on.	1 2 3 4 5
1 2 3 4 5	7.) I wish I were a better person.	1 2 3 4 5
1 2 3 4 5	8.) I'm so weak.	1 2 3 4 5
1 2 3 4 5	9.) My life's not going the way I want it to.	1 2 3 4 5
1 2 3 4 5	10.) I'm so disappointed in myself.	1 2 3 4 5
1 2 3 4 5	11.) Nothing feels good anymore.	1 2 3 4 5
1 2 3 4 5	12.) I can't stand this anymore.	1 2 3 4 5
1 2 3 4 5	13.) I can't get started.	1 2 3 4 5
1 2 3 4 5	14.) What's wrong with me?	1 2 3 4 5
1 2 3 4 5	15.) I wish I were somewhere else.	1 2 3 4 5
1 2 3 4 5	16.) I can't get things together.	1 2 3 4 5
1 2 3 4 5	17.) I hate myself.	1 2 3 4 5
1 2 3 4 5	18.) I'm worthless.	1 2 3 4 5
1 2 3 4 5	19.) Wish I could just disappear.	1 2 3 4 5
1 2 3 4 5	20.) What's the matter with me?	1 2 3 4 5
1 2 3 4 5	21.) I'm a loser.	1 2 3 4 5
1 2 3 4 5	22.) My life is a mess.	1 2 3 4 5
1 2 3 4 5	23.) I'm a failure.	1 2 3 4 5
1 2 3 4 5	24.) I'll never make it.	1 2 3 4 5
1 2 3 4 5	25.) I feel so hopeless.	1 2 3 4 5
1 2 3 4 5	26.) Something has to change.	1 2 3 4 5
1 2 3 4 5	27.) There must be something wrong with me.	1 2 3 4 5
1 2 3 4 5	28.) My future is bleak.	1 2 3 4 5
1 2 3 4 5	29.) It's just not worth it.	1 2 3 4 5