The Effects of a Voice Hearing Simulation on Implicit Fear of Voices

Ciara McEnteggart[1], Yvonne Barnes-Holmes[1] and Funso Adekuoroye[2]

[1]Department of Experimental-Clinical and Health Psychology, Ghent University, Belgium

[2]Department of Psychology, National University of Ireland Maynooth

Corresponding author: ciara.mcenteggart@ugent.be

Abstract

The current study investigated potential changes in implicit negativity of hearing voices in a non-voice hearing student population (*N*=28) subjected to a hearing voices simulation using the Implicit Relational Assessment Procedure (IRAP). On the Baseline IRAP, participants were required to pair voices-as-positive and voices-as-negative statements on alternating trial blocks. Participants were subsequently exposed to a simulation procedure and a Post-Simulation IRAP. At baseline and post-simulation, hearing voices was implicitly evaluated as both positive *and* fearful, however positivity toward voices reduced and negativity increased after the simulation. Interestingly, implicit changes also appeared to be influenced by high delusional ideation.

Keywords: hearing voices simulation, IRAP, implicit measures

Given the breadth of the label of psychosis, and even schizophrenia, researchers have begun to investigate key features that may be specific to these patterns of suffering, specifically hearing voices. This impetus is likely due to two related facts: 1. It is now established that voices are a very commonly reported symptom, not only in diagnoses of psychosis and other psychiatric diagnoses (see Sartorius et al., 1986; Slotema et al., 2012), but also in non-clinical contexts (e.g., Beavan, Read, & Cartwright, 2011); and 2. Social movements, such as the Hearing Voices Movement (aim to normalize and promote the acceptance of unusual experiences) have grown rapidly and are now powerful advocates for social change (Bentall, 2004; Corstens, Longden, McCarthy-Jones, Waddingham, & Thomas, 2014).

Given the prevalence of voice hearing in clinical and non-clinical contexts, and the growing desire for a social change in attitudes toward mental health difficulties, an increasing number of studies have examined attitudes toward voice hearing. Indeed, a vast literature exists demonstrating the presence of negative professional attitudes toward psychological suffering (Schulze, 2007), thus many studies on voice hearing contain interventions that attempt to target these negative attitudes in mental health professionals, often with the aim of targeting stigma, empathy, etc. regarding voices. Many of these studies have included simulations of distressing (or critical) voices within these interventions, due to the prevalence of these types of voices, as reported by voice hearers (Larøi et al., 2012). Overall, these interventions have been associated with positive outcomes and voice simulations have been shown to reduce stigma, but improve empathy, behavioral intentions, and positive attitudes toward voice hearers (Bunn & Terpstra, 2009; Chaffin & Adams, 2013; Dearing & Steadman, 2009; Deegan, 1996; Hojat et al., 2001; Kalyanaraman, Penn, Ivory, & Judge, 2010; Kidd, Tusaie, Morgan, Preebe, & Garrett, 2015; Sideras, Mckenzie, Noone, Dieckmann, & Allen, 2015; Ward, 2015; Wieland, Levine, & Smith, 2015; Wilson et al., 2009).

However, some voice hearing simulation studies have shown less favorable outcomes. For example, Brown, Evans, Espenschade, and O'Connor (2010) found *increased* negative attitudes and an increased desire for social distance (see also Kalyanaraman et al., 2010). Moreover, Brown (2010) reported decreases in willingness to interact with voice hearers, and stronger attitudes centered on help seeking. Interestingly, the mixed findings may pertain to the types of assessment measures researchers have employed. That is, qualitative measures generally produce positive outcomes, while quantitative measures have been more associated with the negative outcomes (Ando, Clement, Barley, & Thornicroft, 2011). Furthermore, the simulation procedures employed vary considerably in presentation (i.e., some were audio simulations and others were virtual reality), length (from 4 to 45 minutes), and content, all of which may also account for the mixed outcomes.

The fact that negative findings have more readily been associated with self-report measures, and that these too have varied considerably across studies of voice hearing simulations, may also speak to the reliance in those studies on a single type of measure. And problems with using only explicit self-report measures are well established (Nisbett & Wilson, 1977; Paulhus, 2002; Wilson & Dunn, 2004). One solution to this situation has seen an increasing number of researchers complement explicit measures with implicit measures, such as the Implicit Association Test (IAT) and the Implicit Relational Assessment Procedure (IRAP), but to date there are no published studies using implicit measures in the context of voice hearing.

The IRAP

The Implicit Relational Assessment Procedure (IRAP) is an automated reaction-time based measure developed specifically from Relational Frame Theory (RFT, Hayes, Barnes-Holmes, & Roche, 2001). It requires participants to pair words and/or pictures, and its basic assumptions are 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 standardized 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). And as the body of supporting evidence for use of the IRAP grows steadily (there are now over 50 published empirical articles), it has come to be used increasingly, and with robust effects, in the study of clinical phenomena (see Vahey, Nicholson, and Barnes-Holmes, 2015).

The Current Study

The current study sought to decrease negativity in terms of fear of voices using a newly-developed, brief, audio voice hearing simulation and the IRAP. The IRAP juxtaposed hearing voices with seeing things. It must be emphasized 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. Notably, the data from the seeing things trial-types were not analyzed because no measure of visual hallucinations was included in the study to control for experience of seeing things. In such an exploratory study, and given the mixed outcomes from simulations noted above, it was difficult to predict whether any change would occur at the implicit level from baseline to post-simulation, however, we hypothesized that we would observe a reduction in negativity at post-simulation.

Method

Setting

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 non-voice hearing participants who were identified as such using current screening methods from a general sample of undergraduate students recruited from the National University of Ireland Maynooth. There were 46 participants, 24 were male and 22 female, with an age range of 18 to 28 years and a mean age of 19.72 years (standard deviation was 1.81 years).

Materials

Self-report measures. Three broad categories of self-report measures were administered. The first series of measures assessed voice hearing and delusional ideation (CAPE). The second set assessed general psychological well-being (AAQ, ATQ and the DASS). The third measured stigma toward mental health difficulties (SAB).

Community Assessment of Psychic Experience (CAPE; Stefanis et al., 2002). The CAPE is a 42-item measure of delusional ideation in the general population (derived from the Peters Delusions Inventory, PDI, Peters, Joseph, & Garety, 1999). This scale has demonstrated adequate reliability: positive dimension (alpha =0.63), negative dimension (alpha =0.64), and depressive dimension (alpha =0.62), and good validity (Konings, Bak, Hanssen, Van Os, & Krabbendam, 2006).

Acceptance and Action Questionnaire II (AAQ-II; Bond et al., 2011). The AAQ-II is a 10-item measure of acceptance of negative private events. This scale has demonstrated adequate internal consistency (0.78 to 0.88), test-retest reliability (0.81 and 0.79), and also demonstrated good construct, concurrent, and predictive validity across several samples (Bond et al.).

Automatic Thoughts Questionnaire (ATQ; Hollon & Kendall, 1980). The ATQ is a 30-item measure of the frequency and believability of negative thoughts. This scale has demonstrated excellent internal consistency with an alpha coefficient of 0.97 and has demonstrated good concurrent validity (Hollon & Kendall).

Depression Anxiety and Stress Scales (DASS–21; Lovibond & Lovibond, 1995). This 21-item DASS comprises three subscales that measure depression, anxiety and stress. 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), and has demonstrated good convergent and discriminant validity, and adequate construct validity (Henry & Crawford, 2005).

Stigmatizing 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. This scale has demonstrated adequate internal consistency (alpha =0.78; Masuda et al.).

The IRAP. The current study involved a Fear IRAP that assessed fearful or normative evaluations of voices. The IRAP contrasted hearing voices with seeing things, using the category labels "hearing voices is" and "seeing things is". Each trial-type presented one of the two category labels. These were accompanied by one of three positive (e.g., "fine") or three negative target stimuli (e.g., "scary"). The positive stimuli were specifically selected to be more ambivalent than highly positive because it seemed unlikely that the non-voice hearing participants would ever have had these types of auditory or visual experiences. 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 1.

INSERT TABLE 1 HERE

Voice hearing simulation. The voice hearing simulation was a newly-developed, brief, audio simulation that comprised two phases in which different topographies of voices heard were presented. Similar to Dearing and Steadman (2009), the first phase involved whispers, and intrusive words or messages (e.g., "You are so stupid"), and the second phase involved arbitrary sounds. Both phases of the simulation were produced by a voice hearer and were a simulation of their own heard voices. The overall simulation lasted 2 minutes, one minute per phase. The inclusion of both phases aimed to increase exposure to different types of voices. increase empathy toward voice hearers, and potentially reduce negativity toward voice

hearing.

Procedure

The current study employed a repeated measures design across five stages: Stage 1, all participants were screened for any experience of *voice hearing* using Item No. 33 of the CAPE; Stage 2, the baseline IRAP; Stage 3, voice hearing simulation; Stage 4, post-simulation IRAP and; Stage 5, self-report measures. Those who screened positively did not participate further (*N*=3), while all remaining participants were then provided with a written explanation of the phenomenon of voice hearing in order to familiarize them with this phenomenon.

Stage 1: Screening. All participants were screened for the presence of voice hearing, using Item No. 33 of the CAPE. All voice hearing participants were excluded from all further stages of the experiment.

Stage 2: 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 consistent with published IRAP studies. For illustrative purposes, see Figure 1 for a schematic representation of the screen presentation of the IRAP.

INSERT FIGURE 1 HERE

Stage 3: 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 4: Post-simulation IRAP. After the voice hearing simulation, participants completed the second exposure to the Fear IRAP.

Stage 5: Self-report measures. Participants completed the five self-report measures in a pre-determined sequence (CAPE, AAQ, ATQ, DASS, and SAB).

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 excluded from the analyses. Mixed between within ANOVAs were conducted for each IRAP exposure (baseline and post-simulation) and a repeated measures ANOVA investigated potential effects of simulation on IRAP scores. Exploratory analyses using the CAPE investigated the potential influence of delusional ideation 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.

Self-Report Measures Data

The means and standard deviations were calculated from each participant's responses on each of the self-report measures and data are summarized in Table 2.

INSERT TABLE 2 HERE

On the AAQ, means revealed that participants overall had low rates of avoidance. On the CAPE, participants had low to moderate rates of delusional ideation. On the DASS, participants had low rates of depression, anxiety and stress. And on the SAB, stigmatizing attitudes toward mental illness were low.

IRAP Data

Scoring of the IRAP followed the standardized approach for transforming latency data into D_{IRAP} scores (see Nicholson et al., 2012). All data from any participant that 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 28 participants.

The mean D_{IRAP} scores on both IRAPs are presented in Figure 2 (visions trial-types are excluded) and the data from both were largely similar. That is, participants showed pro-voices effects on the Voices-Okay trial-type and anti-voices effects on the Voices-Scary trial-type. However, some modest pre-post differences were observed. Specifically, from baseline to post-simulation, the pro-voices effect on Voices-Okay decreased, while the anti-voices effect on Voices-Scary increased.

INSERT FIGURE 2 HERE

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, μ^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, μ^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, μ^2=1.0) and the main effect for simulation approached significance (df=1, F=2.557, p=0.12, μ^2=0.332). However, post-hoc analyses as two dependent t-tests showed no differences from baseline to post-simulation (all p's>0.05). Although, Voices-Scary approached significance (df=27, t=-1.908, p=.067).

Delusional ideation analyses. The mean D_{IRAP} scores for high and low positive dimension delusional ideation scores are presented in Figure 3. On the baseline IRAP, the high group (i.e., those who reported having had *more* delusional ideation) showed greater pro-voices effects than the low group (those who reported having had *less* delusional

ideation) 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 delusional ideation remained the same from baseline to post-simulation, while effects for participants with high delusional ideation became less pro-voices on Voices-Okay and more anti-voices on Voices-Scary.

INSERT FIGURE 3 HERE

A mixed factorial ANOVA demonstrated a significant main effect only for trial-type (F=33.885, p<0.0001, μ^2=1.0). Post-hoc analyses as four independent t-tests found no significant differences between the two delusional ideation sub-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) and on Voices-Okay, the difference approached significance (df=12, t=-1.806, p=0.09). 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 self-report measures. No significant correlations were found between the

trial-types and the self-reported data (all *p's*>0.05). When split by delusional ideation, 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 delusional ideation, the lower the pro-voices effects.

Discussion

This experiment sought to investigate the malleability of fearful evaluations of hearing voices, using a Fear IRAP and a brief voice hearing simulation. All participants had low self-reported levels of stigmatizing attitudes toward mental health difficulties. Yet, at baseline and post-simulation, hearing voices was implicitly evaluated as both positive *and* negative. Positivity toward voices reduced and negativity increased after the simulation, which was contrary to our hypothesis (we predicted a reduction in negativity at post-simulation).

Interestingly, these changes at post-simulation appear to be influenced by participants' experience of delusional ideation. Specifically, the implicit positivity of participants who were low on delusional ideation (i.e., low CAPE score) changed little as a result of directly experiencing voices through the simulation, while those *higher* on delusional ideation became more fearful and less positive. In this study, the implicit positivity in both delusional ideation "sub-groups" at both time points was significant, and while the implicit negativity in those low on delusional ideation group was significant at baseline, it was the negativity in those high on delusional ideation that was significant at post-simulation. Taken together, the simulation procedure appears to have shed light on the likely complex relationship between implicit evaluations and individual experiences of delusional ideation. The study's design also speaks to the potential malleability of IRAP effects.

Overall, these findings appear to contrast the positive effects found in the literature where there were reductions in negativity 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 to complement those studies that have found negative simulation outcomes (e.g., Brown et al., 2010; Kalyanaraman et al., 2010). The negative effects observed here may be explained by the inclusion of negative content in the simulation procedure (i.e., the voices were commanding). That is, while exposure to distressing voices during a simulation procedure may increase empathy for voice hearers as suggested by previous research, it may not decrease negativity toward the hearing of voices, at least at the implicit level. As such, it may be concluded that the use of a *brief* simulation procedure with negative voices does not constitute an effective intervention for *reducing* negative attitudes toward voice hearing (our data showed that this intervention actually serves to increase negativity).

There are a number of variables that may have influenced the current set of results, and which may guide future research in this domain: 1. The duration of the simulation was not manipulated. Future studies could explore whether a longer simulation would produce the same implicit effects. 2. The sample comprised a high proportion of psychology undergraduates exposed to some level of psychological training, which may account for some of the existing positivity. It would be interesting to replicate this study in a sample with no training in psychology. 3. Although all participants were low on stigma at baseline, this was not measured at post-simulation. Future studies could measure stigma at post-simulation in order to investigate if implicit changes are predictive of stigma, and to investigate if these two dimensions are functionally independent. In addition, the finding that those with higher delusional ideation showed the greatest increase in negativity suggests that this may be a critical variable to consider in future research.

Conclusion

The key target variables of simulations should be carefully considered and examined in order to support their clinical utility and to serve a clear purpose in the service of the care of the voice hearer.

Acknowledgements

The data for the current manuscript was collected at the National University of Ireland, Maynooth, and was prepared with the support of the FWO Type I Odysseus Programme at Gent University, Belgium.

References

Ando, S., Clement, S., Barley, E. A., & Thornicroft, G. (2011). The simulation of hallucinations to reduce the stigma of schizophrenia: A systematic review. *Schizophrenia Research, 133* (1–3), 8-16.

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.

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.

Bentall, R. P. (2004). *Madness Explained: Psychosis and Human Nature.* London, England: Penguin Books Ltd.

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.

Bunn, W., & Terpstra, J. (2009). Cultivating empathy for the mentally ill using simulated auditory hallucinations. *Acadamic Psychiatry, 33,* 457–460.

Brown, S. A. (2010). Implementing a brief hallucination simulation as a mental illness stigma reduction strategy. *Community Mental Health Journal, 46*(5), 500-504.

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. *Amrerican Journal of Drug and Alcohol Abuse, 38*(2), 146-154.

Chaffin, A. J., & Adams, C. (2013). Creating Empathy Through Use of a Hearing Voices Simulation. *Clinical Simulation in Nursing, 9*(8), 293-304.

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.

Dearing, K. S., & Steadman, S. (2009). Enhancing intellectual empathy: The lived experience of voice simulation. *Perspectives in Psychiatric Care, 45*(3),173–182.

Deegan, P. (1996). Hearing Voices that are Distressing: A Training and Simulated Experience. National Empowerment Center.

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.

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.

Hojat, M., Mangione, S., Nasca, T. J., Cohen, M. J., Gonnella, J. S., … , Magee, M. (2001). The Jefferson Scale of Physician Empathy: development and preliminary psychometric data. *Educational Psychological Measures, 61*, 349–365.

Hollon, S., & Kendall, P. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research, 4*(4), 383–395.

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.

Konings, M., Bak, M., Hanssen, M., Van Os, J., & Krabbendam, L. (2006). Validity and reliability of the CAPE: A self-report instrument for the measurement of psychotic experiences in the general population. *Acta psychiatrica Scandinavica, 114*(1), 55–61.

Larøi, F., Sommer, I. E., Blom, J. D., Fernyhough, C., ffytche, D. H., … , Waters, F. (2012). The characteristic features of auditory verbal hallucinations in clinical and nonclinical groups: state-of-the-art overview and future directions. *Schizophrenia Bulletin, 38*(4), 724-733.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales* (2nd ed.). Sydney: The Psychology Foundation of Australia.

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.

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. (2012). 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.

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.

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.

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.

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.

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.

Slotema, C. W., Daalman, K., Blom, J. D., Diederen, 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.

Ward, T. D. (2015). Do you hear what I hear? the impact of a hearing voices simulation on affective domain attributes in nursing students. *Nursing Education Perspectives, 36*(5), 329-331.

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.

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.

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.

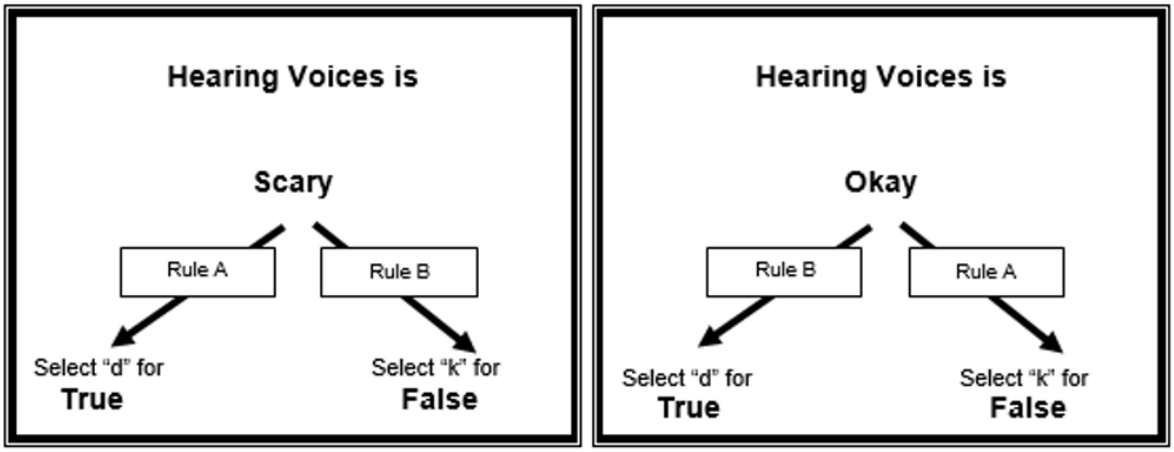

Figure 1. Schematic representation of the two hearing voices trial-types presented in the Fear IRAP. The arrows and text boxes did not appear on the participant's screen, they indicate the correct responses for Rule A and Rule B blocks of trials. **The labels used for the two trial-types are as follows: Voices-Scary (left), Voices-Okay (right).**

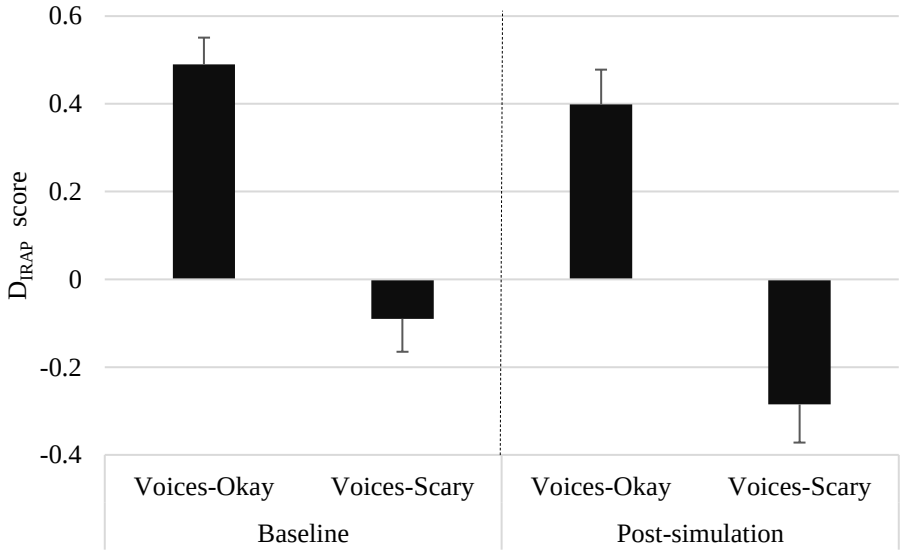

Figure 2. Mean D_{IRAP} scores at baseline and post-simulation on the Fear IRAP. **Positive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

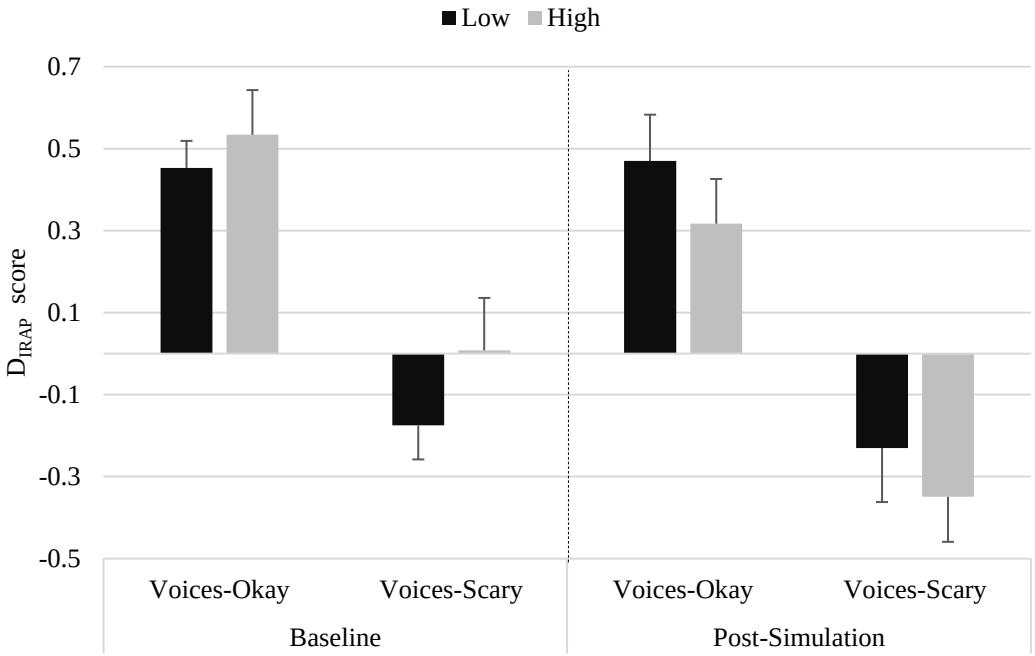

Figure 3. Mean D_{IRAP} scores at baseline and post-simulation on the Fear IRAP for high and low CAPE groups scores. P**ositive D_{IRAP} scores indicate pro-voices effects and negative D_{IRAP} scores indicate anti-voices effects.**

Table 1

Stimuli and Response Options of the Fear IRAP

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

Table 2

Self-Report Data Summary

Scales	
	Mean (SD)
AAQ	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 = 70; CAPE dimensions = 6; DASS total = 126; DASS subscales = 42; SAB = 56.