make [REAL]

Cultivating mental health advocacy and empathy: A longitudinal comparison of the effects of traditional online learning versus interactive perspective-taking

Lynda Joy Gerry[1,2*], Sophie Costin[2], Tim Powell-Jones[2]

[1] Auckland Bioengineering Institute, The University of Auckland, Auckland, New Zealand

[2] Make Real Ltd, Brighton, United Kingdom

* Correspondence:
Corresponding Author
lynda.gerry@makereal.co.uk

Abstract

This paper presents an empirical evaluation of *Hear to Listen*, an interactive online training course aimed to improve mental health literacy, empathic communication, and advocacy behaviors around mental health. We compare two different versions of the course with identical content and duration while testing the efficacy of both and whether the more advanced version might be more effective. Crucially, we were interested in the role of interactive perspective taking (IPT) on our outcomes. We used a traditional e-learning (TEL) version of the app with many features and activities to train mental health literacy. The key difference between the two versions of the application is that the IPT version allowed users to record video and audio responses and then playback the video-based conversation to see and hear themselves from the other person's point of view. Results indicate that both versions of the training led to decreases in biases and increases in cognitive empathy. These results were sustained over a 6-week period. However, the experimental group also exhibited strong increases in advocacy and greater increases in all outcome measures than the TEL group. Moreover, those in the IPT who were lower in trait empathy measures exhibited higher gains through the training than those in the TEL group. Overall, our results indicate that greater self-involvement on the part of the user and incorporation of perspective-taking activities might positively contribute to greater empathic gains and advocacy behaviors.

1   INTRODUCTION

It is estimated that 1 in 6 working-age adults in England experience symptoms associated with poor mental health at least once in their lifetime (McManus et al., 2016). Poor mental health in the workforce can lead to absence, turnover, burnout, and exhaustion, while also reducing productivity and performance (Moll et al., 2015). Mental health issues in the workplace are estimated to cost UK employees around £33 and £42 billion annually (Stevenson & Farmer, 2017). One way to combat these challenges is to improve mental health literacy (MHL) in the workforce. This involves knowing how to recognize the signs and symptoms of mental

health problems and crises and how to respond appropriately (Jorm et al., 1997). One such MHL intervention, Mental Health First Aid (MHFA) is an international training program aimed to improve these interpersonal skills, specifically by changing attitudes towards mental health (decreasing negative attitudes) and increasing motivations to help and intervene (Kitchener & Jorm, 2008). While MHFA has been reported to improve knowledge, attitudes, and behaviours towards mental health problems (Morgan et al., 2018), there is little evidence regarding the impact of MHFA in these contexts (Narayanasamy et al., 2020). Therefore, the first aim of this empirical study was to develop an evidence-based approach to evaluating the efficacy of a MHL intervention specifically designed to increase empathy, decrease stigma, and improve advocacy and support around mental health.

One key aspect to MHL is developing empathy through soft skills training. Soft skills training is a term developed to contrast hard skills such as critical thinking and technical skills to refer instead to human-centered skills such as empathy, interpersonal understanding, and effective communication. Empathy is an especially difficult soft skill to train because there are many barriers to empathy, such as cognitive load, proximity/familiarity, stress/burden, time/energy, self-criticism or lack of self-compassion, difficulties with emotion regulation, and compassion fatigue (for a thorough review, see Fernando & Consedine, 2014). Therefore, we wanted to see if we could enhance cognitive and emotional empathy but also see improvements in real-world behaviors. Developing long-term empathy with real-world behavioral effects is a key obstacle in soft skills training. Specifically, obtaining data that reflects how skills learned are brought into day-to-day life and workplace interactions can be quite challenging. Recently, Herrera and colleagues working on the "Empathy at Scale" project at Stanford University developed a robust empirical methodology for demonstrating long-term efficacy for empathy training interventions (Fernando & Consedine, 2014). Our methodology is adapted from Herrera and colleagues with some adjustments to fit mental illness rather than homelessness.

A current trend in soft skills trainings is to use technology to administer trainings, such as online or e-learning (Ojanperä, 2021), video games ("serious games", Daoudi et al., 2021), and virtual reality (Eckert & Mower, 2020). These learning technologies have been effective in reducing stereotypes and biases (Hirsh et al., 2019; Peck et al., 2013), promoting prosocial helping behaviors (Ahn et al., 2013), and increasing reflection on one's own behaviors (Foster et al., 2016). Hence we wanted to leverage these and similar learning technologies to help administer effective MHL and empathy training. We were specifically interested in helping individuals to develop skills around handling difficult conversations around mental health in the instance of a mental health crisis experienced by a colleague. Therefore, we utilized new learning technologies that offer valuable design techniques to elicit empathy and improve advocacy behaviors which are described in the following section.

1.1 Interactive Perspective-Taking

The design framework that we used in this experiment is categorized as Interactive Perspective Taking. This framework combines key design techniques in new media and emerging technologies to enhance learning soft skills. It combines aspects of participatory learning and perspective-taking pedagogies that involve design features such as simulation, self-feedback, and self-reflection. Participatory learning involves presenting learning content in a way that involves active participation on the part of the user, where the user can interact with the content and receive real-time feedback. This involves more effort on the part of the participant than more passive information presentation styles. Interestingly, greater cognitive effort is also linked to better empathic outcomes, and lower effort has specifically been highlighted as a key limiting factor with newer VR empathy facilitators (Martingano et al., 2021). In this section, we explain each of these design techniques and their components along with empirical support for their efficacy in improving empathy and prosocial behaviors.

Perspective-taking in empathy research is regarded as the ability to imaginatively project oneself into another's viewpoint and to also imagine seeing oneself from another's perspective (Davis, 1994). One design technique to evoke perspective-taking is simulation. *Simulation* involves human actors or digital avatars representing the empathic target (the person for whom participants' empathy is developed) so that participants can practice developing their empathic communication. For example, it has been used in medical training to present clinical scenarios so that medical students, doctors, and nurses can practice communicating with patients. Simulation offers practicing communication skills in risk-free contexts to train reflection (Foster et al., 2016). Simulation trainings in medical education have effectively increased knowledge and confidence (Zigman et al., 2013) and improved attitudes towards patients of varying demographics (Goldfarb & Gorrindo, 2013). Maindonald and colleagues (2020) used a mental health communication skills simulation with non-mental health professionals working in urgent care settings. Similarly, Shao and colleagues (2018) used a simulation-based empathic communication training with neonatal nurses and reported outcomes of a stronger ability to recognize and respond with empathy to patient's emotions. This training simulation led to improvements in confidence, mental health knowledge, and satisfaction with the care received by people experiencing a mental crisis and caretakers. Simulation training involves reflective questions and practical, interactive experiences that motivate learners to develop empathy. Therefore, it targets behavioral component of interpersonal understanding and empathy.

Another perspective-taking design technique is embodied experiences. *Embodied Experiences* in new media design is a technique that places the participant in the embodied perspective of another person. For example, white study participants who embody a black avatar in Virtual Reality (VR) exhibit decreases in racist attitudes (Peck et al., 2013), an effect that has been demonstrated to be sustained over a couple of weeks (Banakou et al., 2016). In another study, male perpetrators of domestic violence showed greater recognition of fear in the faces of their female victims after having an embodied, perspective-taking experience of becoming

their victims in VR (Seinfeld et al., 2018). And finally, participants who experienced the simulation of a puzzle from the perspective of an individual with achromatopsia (color-blindness) were much more likely than a control group to volunteer to help a confederate posing as a person with colorblindness and spent more time with the confederate (Ahn et al., 2013).

In 2008, Petkova and Ehrsson created a variation of an embodied experience that is called a body swap (Petkova & Ehrsson, 2008). A **body swap** is an experience in which participants to experience themselves from the point of view of a partner. The concept of the body swap introduced the notion of potentially re-playing an interaction from another person's point of view. This introduces a new layer of self-feedback into the embodied experience. *Self-feedback* allows participants to observe their own behavior from the point of view of the empathic target to promote reflection on their own behaviors to improve their communication. For example, Foster and colleagues (2016) found that providing medical students immediate feedback about empathic communication during a simulated conversation with a patient expressing a suicide attempt expressed more verbal empathy and were more proficient in suicide risk assessments with future patients. Their simulation training intervention used narrative video vignettes to provide backstory on the patient and offered potential response alternatives during a long conversation with the patient (Foster et al., 2016).

In this study, we integrate these design techniques into our Interactive Perspective Taking model of the *Hear to Listen* course. The *Hear to Listen* course was created by Make Real Ltd in collaboration with Lloyds Banking Group to support confidence, competency, and communication skills related to handling mental health crises in the workplace. Its goal was to improve advocacy, support, empathy, and communication in this challenging situation.

2 METHOD

2.1 Participants

Participants were recruited as volunteers through our partner Lloyds Banking Group. Hence 71 British adult employees of Lloyds Banking Group were recruited for this study. 16 participants dropped out after only completing the pretest and were therefore excluded from the analysis. The final sample (N = 55) consisted of 20 men and 35 women. The ages ranged between 23 and 60 (M = 43.50, SD = 9.10). The majority of the participants were in their mid-late 40s and early 50s (accounts for 50% of the participants). The participants all reside within the UK.

2.2 Design and Procedure

Participants were randomly assigned to one of two conditions, either the Interactive Perspective Taking (IPT; $n = 28$) or the Traditional E-Learning (TEL; $n = 27$) condition. In both conditions, participants received written information about the study and provided informed consent. All participants completed background questionnaires which included demographic questions, the Interpersonal Reactivity Index (IRI) and the Beliefs

about Empathy scale (BAS). The purpose of these scales was to confirm that dispositional empathy and beliefs about its trainability were balanced across the randomly assigned groups.

After completing the background questionnaires, participants completed pre-test questionnaires (Time 0). These included the Inclusion of the Other in the Self (IOS) scale, the 9-item Acquisition Questionnaire (AQ-9), the Reported and Intended Behavior Scale (RIBS), and the Implicit Association Test (IAT). Participants then completed wither the IPT or the TEL task.

Approximately one day after the intervention (Time 1), participants in both conditions performed a mental health training task followed by a post-test questionnaire. This was spaced apart from the pre-test due to cognitive load and time demands so that the measures taken at each time-point would be more balanced. The intervention involved completing the *Hear to Listen* course on Moodle. However, each group accessed a slightly different version of the same course. In the TEL task, participants completed the Moodle course with videos of Isaac and reflections but they did not record or playback their own video. In the IPT task, participants recorded their video response to Isaac and then the videos would be played back from Issacs's point of view. The information, task sequence, and course content was exactly the same in both tasks, and both tasks lasted approximately 30-45 minutes.

Immediately after the intervention, participants completed a post-test questionnaire. During the post-test (Time 1), participants completed the IOS, RIBS, the Self-Stigma of Mental Illness Scale (SSMI), Proposition A, and the IAT.

Two weeks later, participants completed the first follow-up (Time 2). Here they complete the AQ-9 (adjusted to be Beth, a new employee at work), RAS, RIBS, and a Letter Writing Task (Writing to an Elected Official).

Four weeks after the intervention at Time 1, participants completed a second follow-up (Time 3). During this follow-up, they completed the SSMIS, RIBS, and the Town Hall Invite.

Six weeks after the intervention at Time 1, participants completed a third and final follow-up (Time 4). They completed the Attitudes Towards Mental Health Questionnaire, RAS, RIBS, a Letter Writing Task (Letter to a Friend), and the IAT.

One month after the experiment was complete, participants received full debriefing information and disclosure about which group they were in and the initial results from the empirical study.
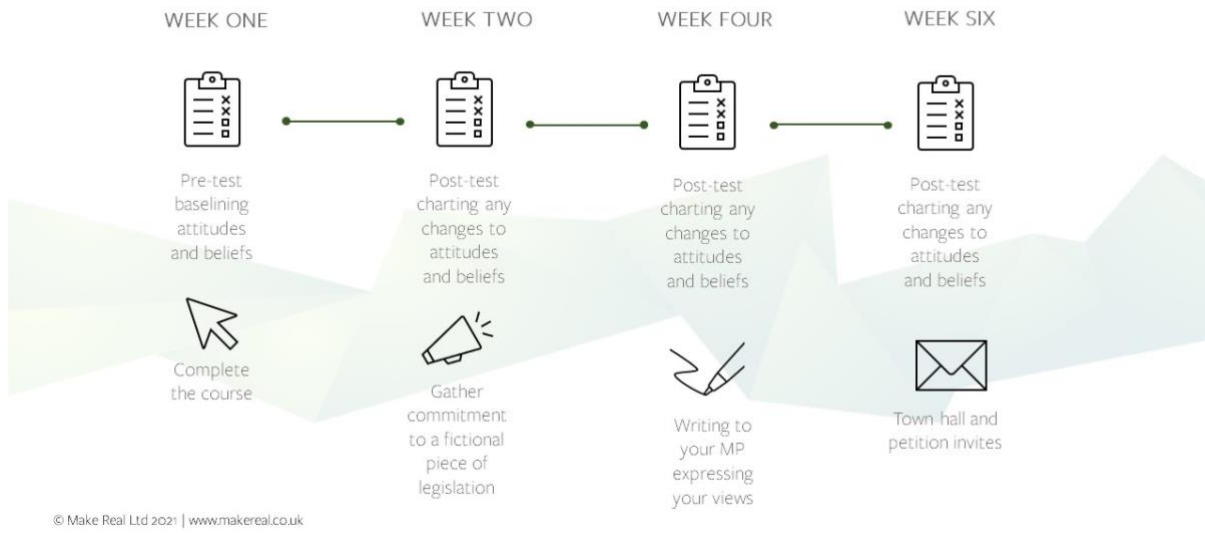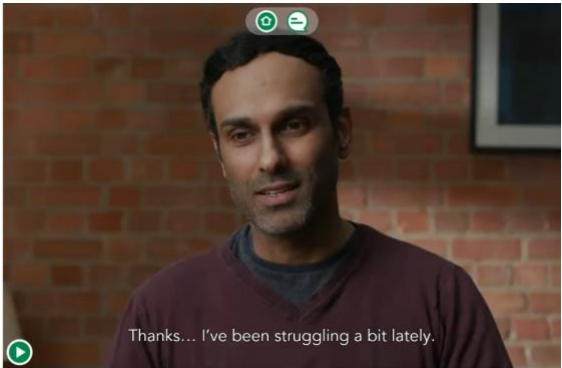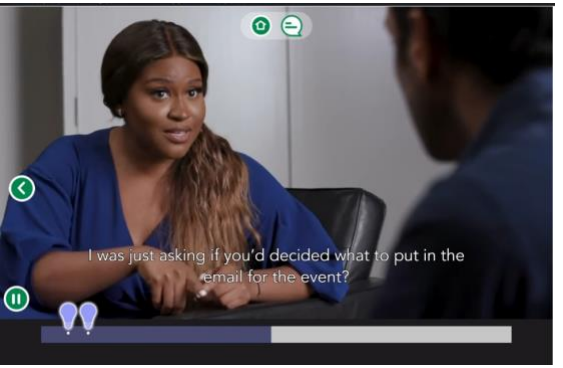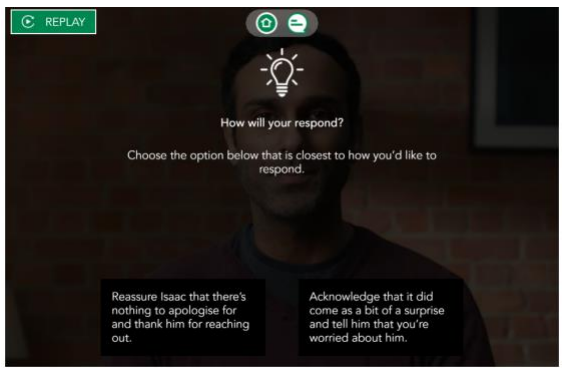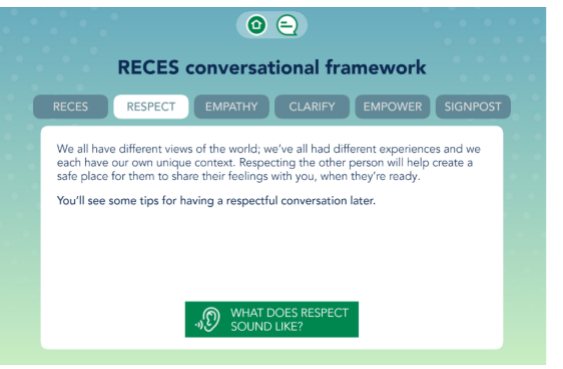
Figure 1. *Overview of experimental procedure.*

2.3 Intervention

The *Hear to Listen* application simulates a sudden video call from a colleague named Isaac who is experiencing a mental health crisis. The application is hosted on the learning management platform Moodle. When the course starts, participants see a series of statistics about mental health and how common mental health issues are. Participants see Isaac calling and can click a button to answer. When they answer, Isaac is clearly distressed and explains that he has been struggling lately and expresses having trouble seeing a point to getting out of bed. The user is then given an opportunity to reflect on how they will respond to Isaac. In the TEL version, users select between two boxes presenting different language for their responses. In the IPT version, users record their own video and audio response to Isaac. Then the user practices responding verbally to Isaac. The video continues, and Isaac is increasingly distressed. Finally, Isaac hears a knock on the door and asks the participant to forge that he said anything. After this first conversation run-through, we ask participants to complete an Empathy and Personal Distress questionnaire rating levels of positive and negative empathic emotions. Then users rate how confident they are in their response. In the IPT version, the user is told that they will have the same conversation again, but from Isaac's point of view and afterwards are asked how they felt about their responses. In both versions, the users then run through an interactive video exercise playing a scene from Isaac's workday on the previous day. During this interactive video, participants are asked to mark whenever they notice a sign or a symptom of Isaac's poor mental health. In both versions of the app, the users click through some information about mental health signs, symptoms, and how to reduce its stigma. Then the users see a RECES model with advice on how to respond to Isaac. This model instructs participants to use language to respect, empathize, clarify, and empower (RECES) someone experiencing a mental health issue. Users also see some screens with some "Dos and Don'ts" for how to respond to someone with a mental health

crisis. Finally, participants have a second conversation with Isaac. This follows the same structure as the first conversation with regards to the differences between the two groups. During this conversation, Isaac provides some background about why he is struggling and the user has the opportunity to practice the RECES model, suggest that he take some time off, and seek mental health support resources at the workplace. Finally, the user completes the Empathy and Personal Distress and confidence check questionnaires once more. Before exiting the app, the user is presented with results of their confidence level changes and prompted to reflect on their journey using the app.

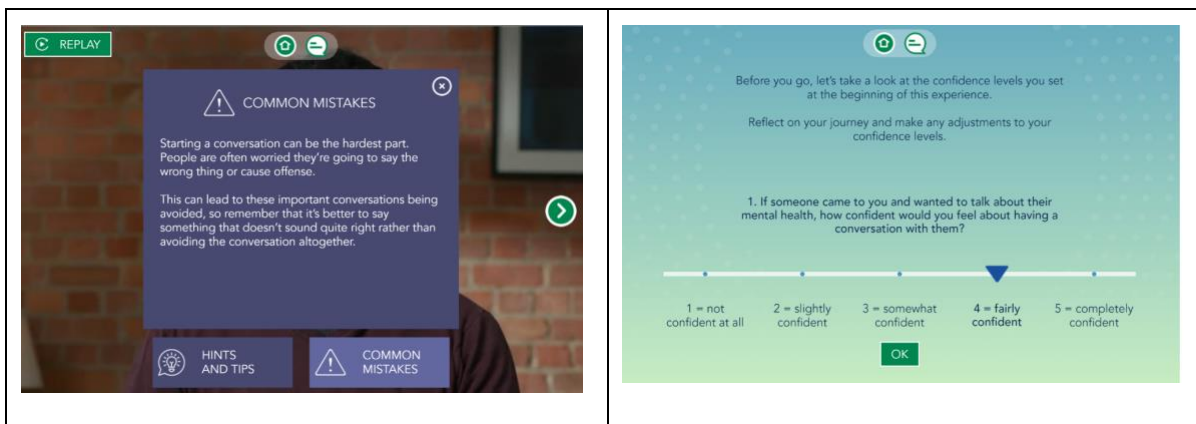| Video	Interactive video
Multiple choice interactions	Text and image content
Feedback	Self-evaluation and reflection activities

Figure 2. *Overview of design elements used in the TEL version of the Hear to Listen application. The TEL version of the application was developed with subject matter experts and follows established best practice for traditional eLearning, and using a combination of the following elements: video, interactive video, multiple choice interactions, text and image content, feedback, and self-evaluation and reflection activities*

The TEL and IPT versions of the Hear to Listen app were identical except that the IPT version contained an additional activity type aimed at encouraging more effortful engagement. These were video-based interactions that required learners to record their response via webcam to certain questions. Once they were happy with what they had recorded they would see themselves from their colleague's point of view and reflect on their own verbal and non-verbal communication.

Record your response	Watch yourself back	Evaluate your response
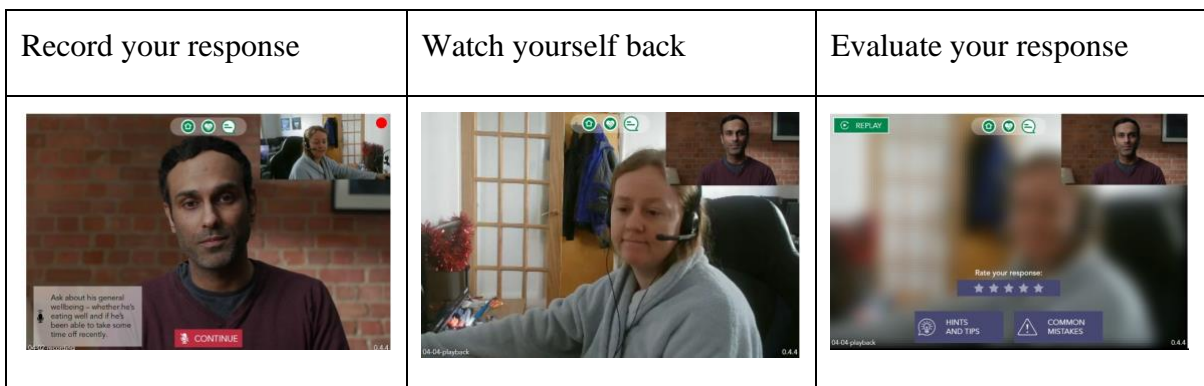
Figure 3. *Figure depicting the additional IPT activity of video-based interactions.*

2.4 Measures

Population Measures

IRI empathy dimensions. The Interpersonal Reactivity Index (IRI) is a 28-item questionnaire in which participants rate their empathic dispositions on a 5-point Likert scale ranging from 0 (doesn't describe me at all) to 4 (describes me very well). Under the classic constructions of the IRI, there are four subscales containing seven items each. These subscales include perspective-taking

(PT), empathic concern (EC), personal distress (PD), and fantasy (FT). For the purposes of this study, the fantasy (FT) subscale was removed. The reason that this subscale was excluded was to reduce the amount of time that participants would be completing the personality and pre-intervention questionnaires, given that the intervention lasted 30-45 minutes. Additionally, this omission was based on previous research from Herrera et al. (2019), who explained that because FT is designed to target individuals' capacity to imaginatively embody fictional characters it is less relevant for a perspective-taking intervention. Therefore, seven items made up each of the remaining three subscales: PT (e.g., "I try to look at everybody's side of a disagreement before I make a decision."), EC (e.g., "I often have tender, concerned feelings for people less fortunate than I."), and PD (e.g., "In emergency situations, I feel apprehensive and ill-at-ease.").

Beliefs about Empathy Scale. The Beliefs about Empathy scale is a 12-item scale answered on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). The scale measures the degree to which people believe that empathy is something that can be learned and/or controlled. Sample items include "example 1 here" and "example two here." The scale contains two different subscales, controllability ($M = 4.39$, SD = 1.02, Cronbach's alpha = .72) and implicit theories ($M = 4.21$, SD = 1.23, Cronbach's alpha = .87). The overall reliability of this scale was good, Cronbach's alpha = .67. This scale is used as a sample check to ensure that there were no significant differences in beliefs across empathy across conditions, as previous research has demonstrated that participants who believe they can control their empathic responses actually exert more empathic effort than those who believe they have no control (Schumann et al., 2014).

Outcome Measures

Inclusion of the Other in the Self. The Inclusion of the Other in the Self (IOS) scale is a measure of closeness and connectedness to a target individual or group. The IOS Scale is a single-item pictorial scale that presents a series of increasingly overlapping Venn diagrams. For this experiment, "a personal with a mental health condition" was used as the target. The prompt for the questionnaire was, "Please indicate how close you feel to a person with a mental health condition, where X indicates a person with a mental health condition. Select the pair of circles with the degree of overlap that best represents your current relationship." The IOS is a 7-point scale with higher numbers corresponding to greater overlap of the circles and indicating a closer relationship (*M* = 4.34, *SD* = 1.44).
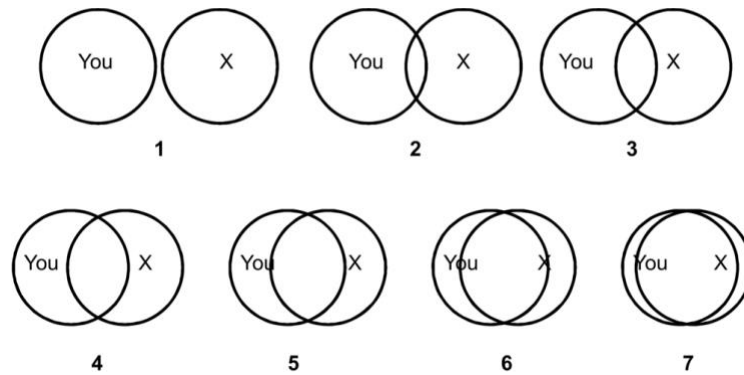
Figure 4. *The Inclusion of the Other in the Self (IOS) scale, where X indicates a person with a mental health condition.*

Empathy. Participants were asked to rate the extent to which they felt four emotions throughout the intervention: softhearted, touched, sympathetic, or compassionate. These items were rated using a 7-point Likert scale (1 = Not at All, 7 = Extremely) and the results created an index of empathic concern. This measure was adapted from . The reliability of the scale was good, Cronbach's alpha = .72 (*M* = 4.82, *SD* = .95).

Personal Distress. Using a 7-point Likert Scale, participants rated the extent to which they felt four emotions throughout the intervention: uneasy, troubled, distressed, or disturbed. The results created an index of personal distress. This measure was also adapted from Batson, Early, and Salvarani (Batson et al., 1997). The overall reliability of the index was strong, Cronbach's alpha = .61 (*M* = 2.97, *SD* = .96).

Negative Attitudes Towards Mental Illness. For this measure we used the Acquisition Questionnaire (AQ-9) in two different versions depicting either Harry or Beth. These are both existing versions of the AQ-9 scale. The AQ-9 is a 9-item scale (shortened from the original AQ-27) that measures attitudes in response to a short narrative vignette about a person with a mental illness. Participants rate their extent of agreement with a series of statements that represent negative biases towards mental health. The Self-Stigma of Mental Illness Scale (SSMI) and Attitudes Towards Mental Health questionnaires were also used in follow-up measures to add variety to the experiment without repeating questionnaires while indexing the same dependent measure.

Positive Attitudes Towards Mental Illness. For this we used an adapted version of the Recovery Assessment Scale (RAS), which indexes positive attitudes regarding the capabilities, empowerment, self-determination, and recovery potential of people who struggle with mental illness. This version of the scale was adapted from Corrigan and colleagues (2014). They found that in addition to decreasing negative stigmas, increasing affirmative attitudes around mental illness is also important for increasing social inclusion and acceptance (Corrigan et al., 2014). The adapted version of the RAS contains a 13-item questionnaire with Likert scale items ranging from 1 "strongly disagree" to 5 "strongly agree." The adapted RAS contains four subscales: Ability to Succeed, Empowerment, Self-Advocacy, and Social Support. The overall reliability of the index was strong, Cronbach's alpha = .67 (*M* = 2.66, *SD* = .39).

The Reported and Intended Behavior Scale. The Reported and Intended Behavior Scale (RIBS) is an eight-item self-reported behavioral measure that investigates reported and intended behaviors towards mentally ill persons among four different contexts: 1) living with, 2) working with, 3) living nearby, and 4) continuing a relationship with someone with a mental health issues (Evans-Lacko et al., 2011). The scale is intended to measure current and intended behavioral discriminations towards people with mental health problems. Participants rank their willingness on a 5-point Likert scale from "Strongly Agree" to "Disagree Strongly." Behavior change can be measured as the extent of increased agreement in the willingness to engage with mentally ill persons in the future across these four contexts from pre-test and post-test, as well as longitudinally. Higher scores indicate a high intended social proximity, which is how close a research participant is willing to get to someone with a mental health issue.

Behavioral Measures

Agreement with Proposition A. Proposition A includes a script of a proposition to increase the National Institute of Health (NIH) spending on Statutory Sick Pay in the UK and increase accessible and affordable mental health services for employees. This proposed was drafted based on real propositions at the time the study was conducted that support improving mental healthcare as sick leave in the UK (Text A in Appendix). After participants read Proposition A, they were asked to rate the extent to which they would be willing to sign it. This was rated using a 5-point Likert scale ($M = 3.94$, $SD = .93$) ranging from 1 "not at all" to 5 "extremely likely". This measure was implemented at Time 0 (immediately following the intervention).

Letter Writing. Participants were asked to write a letter to their elected officials (their member of parliament) regarding the issue of mental health in the workplace at Time 1 (2 weeks after the intervention), and a letter to a friend explaining what they have learned about the issue of mental health at Time 3 (8 weeks after the intervention). The letter writing task was designed to measure civil engagement, emotional states, and advocacy on the issue of mental health. This allowed us to conduct a linguistic analysis of the language used in the letters using 7 categories: word count, positive emotion, negative emotion, social, anxiety, I, and we. These categories were chosen to quantify affect, advocacy, and to understand the extent to which participants included or excluded themselves as part of the solution when writing about the issue of mental health. All linguistic analyses were conducted using Linguistic Inventory and Word Count (LIWC) software.

Town Hall Meeting. As a test of willingness to help, we invited participants to participate in a Town Hall meeting on the topic of mental health. This invite was administered independently from the experiment so that participants did not know it was part of the experiment. This invite was sent after Follow-Up 2. This invite was set up such that it would require participants to sacrifice their personal time (non-work hours) to attend a workshop to learn more about mental health in the workplace.

3   Results

Population Variables

A one-way analysis of variance (ANOVA) showed that there was no significant difference in the participants' trait-levels of empathy on any subscale of the IRI across the two conditions (Perspective Taking: t(53) = -.936, *p* = .177; 95% Confidence Interval (CI) [-2.63, 0.96], Empathic Concern: t(53) = -.220, *p* = .413; 95% Confidence Interval (CI) [-1.74, 2.16], and Personal Distress: t(53) = -1.11, *p* = .136; 95% Confidence Interval (CI) [-3.29, 0.94]. The IRI measures PT, the ability to project oneself in the point of view of others (*M* = 2.98, SD = 0.51, *Cronbach's alpha* = .56), EC, the ability to feel empathic concern for others (*M* = 2.57, *SD* = 0.48, *Cronbach's alpha* = .65), and PD, the experience personal distress in stressful situations with others (*M* = 1.25, SD = 0.56, *Cronbach's alpha* = .64). Additionally, participant's levels of perspective-taking (PT) were highly correlated with their levels of empathic concern (r(53) = .34, p=.01), and significantly negatively correlated with personal distress (r(53) = -.28, p=.04). That is, those who were better at perspective-taking were also strong in empathic concern but experienced less personal distress.

Descriptive Statistics of IRI Subscale Scores and Total IRI Scores

Scale/Sub-scale	Mean	SD	α
Perspective-taking	2.57	0.48	0.56
Empathic Concern	2.98	0.51	0.65
Personal Distress	1.25	0.56	0.64
IRI Total Score	47.47	6.12	0.58

Table 1. *The maximum possible score is 4 for IRI sub-scales (scores are means of a subject's response to the items on a given sub-scale) and 60 for the IRI total scale. α = Cronbach's alpha.*

There were also no significant differences between the two conditions regarding beliefs about empathy (Controllability: *t*(53) = 0.27, *p* = .621; 95 CI [-.73, .28], Implicit Theories: *t*(53) = -1.12, *p* = .241; 95 CI [-.69, .18]). These results show that there was a balance across conditions in terms of the way that people think about empathy and in the way that they believe they are able to control their empathic responses, showing that random assignment was successful across conditions.

Outcome Variables

IOS. Paired samples t-test was used to compute the significance within-subjects of the pretest-posttest increase in self-reported measures from the two questionnaires. On the IOS scale, the mean score for each participant increased from 3.87 on the pretest to 4.82 on the posttest. Each participants' score significantly increased from pretest-posttest (p < .001). Notably, there was a difference by gender (women were higher) on mean scores on the IOS prettest, but not on the posttest (M = 3.6, SD =

Within-Subjects Changes in IOS Scores (Pre-Post)

		Mean	N	Std. Deviation	Std. Error Mean	Mean Difference	Significance
IOS	Pretest	3.87	55	1.45	0.20		
	Posttest	4.82	55	1.43	0.19	0.95	<.001

Table 2. *Descriptive statistics for the Inclusion of the Self (IOS) scores on pretest and posttest IOS and mean difference (pre-post) for all study participants. Paired samples t-test of significance within-subjects shows that post-test scores were significantly higher than pretest scores.*

Univariate ANOVA was used to compute the significance between subjects. The average difference between pre-test and post-test was 1.46 for the IPT and .407 for the TEL group. There was a statistically significantly higher increase in pretest-postest IOS scores for the IPT group as compared to the TEL group (p=.012). This has an effect size of 0.7, meaning that around 75% of the control group were below the mean of the experimental group.

Between Groups IOS Descriptive Statistics

Dependent Variable: Difference in IOS Scores			
Group	Mean	Std. Deviation	N
TEL	0.4074	1.44806	27
IPT	1.4643	1.55116	28
Total	0.9455	1.58018	55

Table 3. *Descriptive statistics for the Univariate ANOVA computing between-subjects effects for IOS score.*
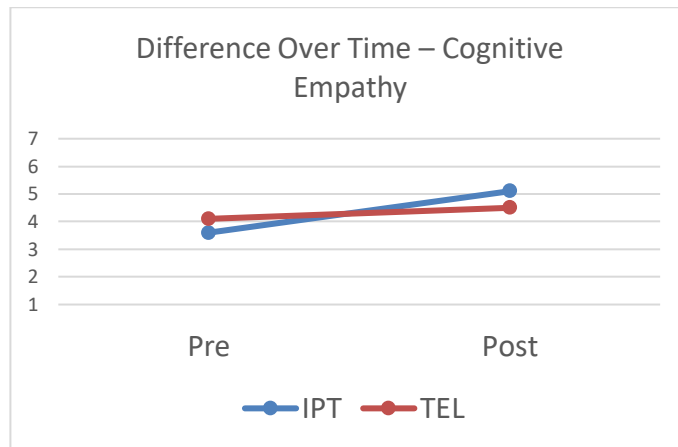

Figure 5. *Chart depicting change over time for IOS, which is a measure of cognitive empathy.*

Participants who were high in the perspective-taking (PT) subscale of the IRI in the TEL group decreased on their IOS score from pre-test to post-test (M = -.169, SD = 0.98). Participants who had high PT

in the IPT group increased their IOS score slightly (M = .034, SD = 1.08). Those in the IPT group who were low on PT had a greater increase on IOS than those who were high in PT (M = .622, SD = 0.33). Interestingly, who were low on PT in the TEL group decreased their IOS score (M = -.456, SD = 1.04), whereas participants who were low in PT in the IPT group increased (M = .621, SD = .336).

High and Low Perspective Taking and IOS Score Change

Group	Perspective Taking	Mean IOS Change	N	Std. Deviation
TEL	Low	-0.456	12	1.04
	High	-0.169	15	0.983
IPT	Low	0.622	12	0.33
	High	0.034	16	1.08

Table 4. *Table depicting the mean difference in IOS scores pre-post by group filtered by high and low perspective-taking on the IRI. This shows that those in the IPT group have a more significant change in their scores, particularly those who are low on PT.*
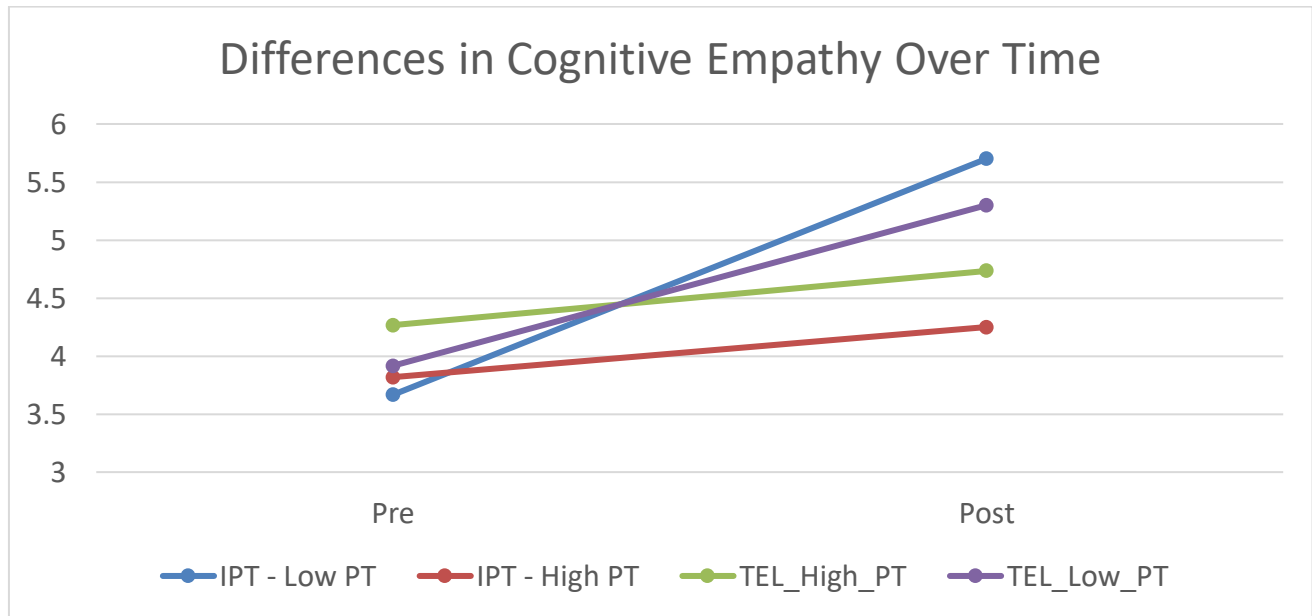
Figure 6. *Figure depicting IOS changes over time as a function of high versus low perspective taking and in each experimental group (TEL and IPT). Hence, lower PT is linked to a greater change in IOS scores. Those who are low in PT in the IPT group have the most significance IOS score change over time.*

Participants who were low on the empathic concern (EC) subscale of the IRI had an increase in their IOS scores for the IPT group (*M* = .361, *SD* = .582), and a decrease for the TEL group (*M* = -.337, *SD* = 1.06). An ANOVA was conducted and revealed that when controlling for empathic concern scores, the difference between the groups is still statistically significant (p=.024).

High and Low Empathic Concern and IOS Score Change

Group	Empathic Concern	Mean IOS Change	N	Std. Deviation
TEL	Low	-0.261	13	0.979
	High	-0.337	14	1.06
IPT	Low	0.361	14	0.581
	High	0.219	14	1.13

Table 5. *Table showing the mean difference in IOS scores pre-post by group filtered by high and low empathic concern on the IRI. This shows a similar pattern as PT. Those who are low on EC have a higher change in their scores. However, the TEL group is decreasing their scores while the IPT group is increasing their scores. Those who are low in EC have slightly higher change in scores.*

Participants who were low on Personal Distress (PD) had increases in their IOS score pre-posttest, but this increase was much greater in the IPT group (M = .278, SD = .867) than the TEL group (M = .020, SD = .786). Those who were high on PD decreased their IOS score in the TEL group (M = -.727, SD = .949) and increased in the IPT group (M = .375, SD = 1.02).

High and Low Personal Distress and IOS Score Change

Group	Personal Distress	Mean IOS Change	N	Std. Deviation
TEL	Low	0.02	14	0.786
	High	-0.728	13	0.95
IPT	Low	0.28	14	0.867
	High	0.375	14	1.02

Table 6. *Table showing the mean differences in IOS pre-post IOS scores by group (TEL and IPT) filtered by high and low persona distress. Results indicate that those who are high in personal distress decreased their IOS scores in the TEL group but increased in the IPT group.*

Empathy and Personal Distress. Paired samples t-test was used to compute pre-post changes in empathy and personal distress. Results indicate that all participants increased in positive empathy (softhearted, touched, sympathetic, and compassionate) and decreased in empathic distress (uneasy, troubled, distressed, and disturbed). The results were especially significant (p<.001) for pre-post decreases in feeling uneasy, troubled, and disturbed.

Within-Subjects Changes in Empathy and Personal Distress

		Mean	Std. Deviation	Mean Difference	Significance
	Pre	4.06	1.25		
Softhearted	Post	4.53	1.79	0.47	0.036

	Pre	4.53	1.33		
Touched	Post	4.95	0.33	0.41	0.075
	Pre	4.23	1.20		
Uneasy	Post	1.88	0.39	-2.35	<.001
	Pre	4.76	1.15		
Troubled	Post	2.29	0.39	-2.47	<.001
	Pre	2.94	1.85		
Distressed	Post	1.53	0.33	-1.41	0.012
	Pre	4.64	1.22		
Sympathetic	Post	5.52	0.17	0.88	0.010
	Pre	3.12	1.83		
Disturbed	Post	1.30	0.31	-1.82	<.001
	Pre	4.76	1.35		
Compassionate	Post	5.56	0.17	0.82	0.019

Table 7. *Table depicting pretest and posttest scores for each of the 8 empathic emotions, showing significant decreases in empathic distress and increases in empathic concern.*

Univariate ANOVA did not reveal any significant differences between groups on the post-test scores. However, Univariate ANOVA of the mean difference in scores (pre-post) indicated statistical significance in "softhearted" ($M = 2.27$ in the IPT group versus $M = 1.05$ in the TEL group, p = .017) and "touched" ($M = 2.49$ in the IPT group versus $M = 1.21$ in the TEL group). The IPT group had a more significant increase in these states than the TEL group. Moreover, compared to the TEL group, the IPT group had a statistically significant decrease in feeling "distressed" ($M = -1.24$ for the IPT group and $M = -.40$ for the TEL group, p = .010) and "troubled" (p = .045).

<u>Negative Attitudes Towards Mental Health.</u> In the AQ-9 questionnaire, each question represents a variable that is correlated to a specific stigma around mental health. These 9 stigma variables include: pity, danger, fear, blame, segregation, help, and avoidance. Paired samples t-tests were performed to measure the

differences in pretest-posttest scores for each item in the questionnaire within subjects. Results reveal statistically significant decreases in pity ($p < .001$), danger ($p < .001$), fear ($p < .001$), blame ($p = .044$), segregation ($p = .032$), avoidance ($p = .002$), and coercion ($p < .001$). Results also show a statistically significant increase in help ($p < .001$), which is scored on the AQ-9 as a negative attribution to mental health. This is because it has a stereotype associated that people with mental health conditions need help. However, the training exercise was focused on increasing helpful and supportive behavior, and therefore this increase indicates that participants were more willing to want to help someone with a mental health condition. Results were not statistically significant for anger ($p = 0.134$). This is due to the fact that the participants' scores were quite low to start ($M = 1.41$), and stayed low ($M = 1.22$ on post-test). A similar pattern occurs for "blame", which had a slightly weaker significance pre-post, but this is because values were low for both pretest and posttest.

Paired Samples t-test for AQ-9 Pre-Posttest Scores

	Correlation	Significance		Mean Difference	Significance	
		One-Sided p	Two-Sided p		One-Sided p	Two-Sided p
Pity	0.59	<.001	<.001	-0.98	<.001	0.002
Danger	0.402	0.003	0.005	-1.87	<.001	<.001
Fear	0.52	<.001	<.001	-1.66	<.001	<.001
Blame	0.339	0.01	0.02	-0.17	0.044	0.088
Segregation	0.021	0.445	0.89	0.57	0.032	0.065
Anger	-0.075	0.311	0.621	-0.20	0.134	0.269
Help	0.387	0.004	0.007	1.45	<.001	<.001
Avoidance	0.367	0.006	0.011	-0.79	0.002	0.004
Coercion	0.463	<.001	0.001	-1.20	<.001	<.001

Table 8. *Paired samples t-test indicating the within-subjects descriptive statistics for the Pre-Post scores on each item (1-9) in the AQ-9 questionnaire and the mean differences in the scores from the pre-test to post-test.*

Univariate Analysis of Variance (ANOVA) was conducted to measure the difference in pretest-posttest scores on each questionnaire item of the AQ-9 between-subjects. The results were not statistically significant for pity ($p=.619$), danger ($p=.785$), fear ($p=.563$), segregation ($p=.229$), anger ($p=.865$), help ($p=.374$), avoidance ($p=.388$), or coercion ($p=.224$). The results were statistically significant for blame ($p=.037$), where the IPT group had a more significant decrease. That is, the IPT group attributed less blame to people with a mental health condition after the training than the TEL group.

Between Subjects Comparison of Blame Score Decreases				
Dependent Variable: Difference_AQ9_4_Blame				
Group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
TEL	-0.045	0.137	-0.322	0.231
IPT	-0.36	0.129	0.1	0.62

Table 9. *Table showing descriptive statistics for between-subjects differences in pre-posttest scores for the item 4 measuring "blame" on the AQ-9.*

Positive Attitudes Towards Mental Illness. Paired samples t-test on the adapted RAS total score and subscales indicated that participants significantly increased their total scores from pre-test to the final follow-up (M = 1.11, SD = 7.64, p = .002). Additionally, all positive attitudes except for social support significantly improved from the pretest to FU3. These include: ability to succeed (M = 1.06, SD = 3.89, p = .004), empowerment (M = .014, SD = 3.5, p < .001), and self-advocacy (M = 0.06, SD = 1.62, p <.001). Interestingly, participants did not significantly change their attitudes that a person with a mental health condition can rely on others within their social support network (M = 0.24, SD = 1.13, p = .277). The item for this subscale was, "People with mental illness have people they can count on." It is possible that the *Hear to Listen* simulation made participants uncertain as to whether or not people with a mental illness have people they can count on to support them since it involved a call out of the blue from a colleague. Alas, the self-advocacy high scores and significant increases indicate that participants feel that a person with a mental health condition knows when to ask for help.

Univariate ANOVA did not produce any statistically significant differences in RAS total scores by group on the FU3 measure as compared to the FU1 measure. "Ability to Succeed" was the only subscale with statistical significance between the two groups (*M* = 21.02, *SD* = 3.14 in the IPT group and *M* = 19.52, *SD* = 2.75 in the TEL group, p = .049).

Paired Samples Statistics for RAS Scores

		Mean	Std. Deviation	Mean Difference	Significance One-Sided p	Two-Sided p
	FU3	46.72	7.297			
Total RAS Score	Pre	45.62	6.91	1.11	**0.002**	0.003
Ability to Succeed	FU3	20.04	2.94			
	Pre	18.98	3.44	1.06	**0.036**	0.073
	FU3	17.06	3.52			
Empowerment	Pre	16.92	3.31	0.14	**<.001**	<.001
	FU3	5.96	1.67			
Self-Advocacy	Pre	5.90	1.33	0.06	**<.001**	<.001
Social Support Network	FU3	3.46	0.89			
	Pre	3.22	0.79	0.24	0.277	0.554

Reported and Intended Behavior Scale. On the Reported and Intended Behavior Scale (RIBS), paired samples t-test revealed an increase in scores within-subjects. That is, participants exhibited a statistically significant increase in their RIBS scores from pretest to posttest (p<.001).

Paired Samples Statistics for Pre-Posttest RIB Scores

		Mean	Std. Deviation	Mean Difference	Std. Deviation	Significance
RIBS	Pre	19.9423	3.53921			
	Post	21.9231	2.8412	-1.98	3.09	<.001

Table 10. *Paired samples t-test descriptive statistics for the means of difference in scores from pretest to posttest. Paired Samples t-test results indicate statistical significance (p<.001) for within-subjects pretest-posttest scores.*

Univariate Analysis of Variance (ANOVA) was conducted to analyze the difference in the mean scores between the two groups. There was a very significant difference in the increase in scores between the two groups (p < .001). That is, the IPT group had a higher increase in their RIBS score as compared to the TEL group.

Behavioral Variables

Letter Writing. The letter writing data was analyzed using the linguistic analysis software Linguistic Inventory and Word Count (LIWC) version 2022-02-18 for Mac. Category analysis focused on tone, affect, social attitudes, and values orientation towards health and particularly mental health. Significance was computed with a one-way ANOVA setting experiment group (IPT and TEL) as the dependent variable and categories set as the independent variables.

In the *Letter to an Elected Official* writing task for Follow-Up 1, the IPT group used language that reflected stronger positive emotions (p=.028). The IPT group also used significantly more references to wellness as a theme in their letter (p=.004) and a stronger moral tone (p=.030). The IPT had a more positive tone than the TEL group, whereas the TEL group had a more negative tone. The IPT group had more language exhibiting a prosocial intent approaching statistical significance (p=.061).
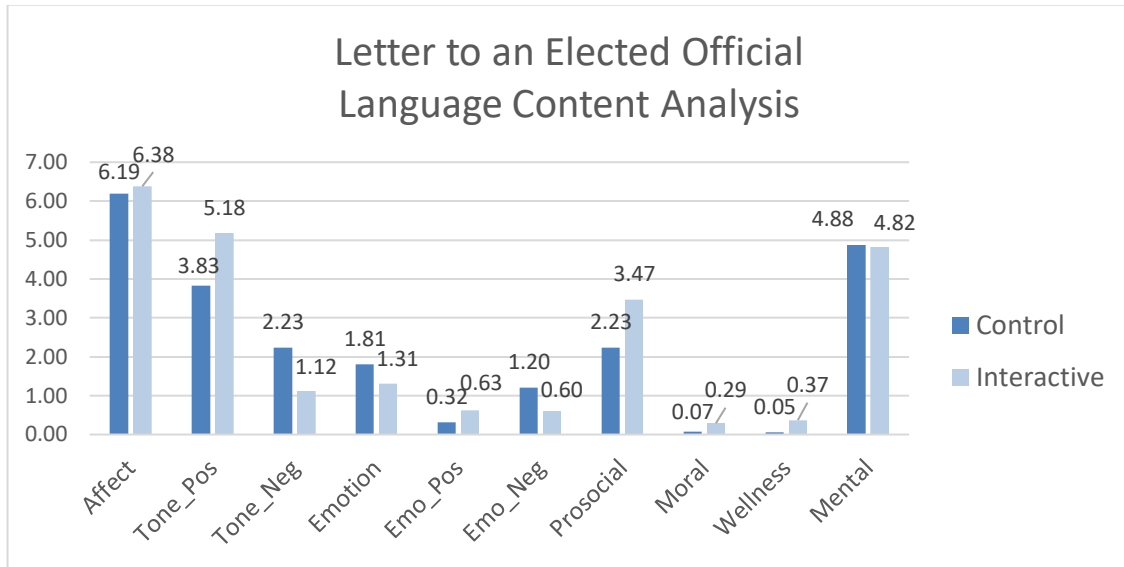
Figure 7. *Graph of means in each LIWC category variable by group (Control is TEL and Interactive is IPT).*

The most common words used in the *Letter to an Elected Official* writing task were mental health, support, people, issues, and community (see Table 1). The IPT group used the word "support" more than the TEL group 82% of participants in the IPT group used this word in their letter, compared to only 65% of participants in the TEL group. 41% of the TEL group participants used the word "awareness" in their letter, whereas no one in the IPT used this word. 45% of participants in the IPT group used the word "services" in their letters, versus only 11% in the TEL group.

Word	Frequency	Rows with Word	Percentage of Rows with Word
health	123	38	95.0
mental	114	37	92.5
support	72	30	75.0
people	57	26	65.0
issues	47	27	67.5
community	32	14	35.0
available	20	11	27.5
services	19	12	30.0
see	19	12	30.0
local	17	11	27.5
needs	15	11	27.5
area	14	10	25.0
work	11	6	15.0
awareness	10	9	22.5
time	9	9	22.5
understanding	9	7	17.5
experience	9	7	17.5
suffering	9	6	15.0
stigma	8	8	20.0

Table 11. *Table summarizing most used words (for both groups) in the Letter Writing Task, as analyzed in the Language Inventory and Word Count (LIWC) software.*

Figure 8. *Word cloud indicating the highest prevalence of words used in the Letter to an Elected Official writing task.*

Proposition A. Univariate ANOVA comparing the scores between the two groups revealed a statistically significant higher support for Proposition A in the IPT group as compared to the TEL group ($p <$ .001). Interestingly, women were statistically in greater support of this petition than men. The mean score for women was 4.33 (out of 5 indicating increasing levels of support), as compared to 3.94 for men [$t(52)=-1.69$, $p=.05$].

Proposition A Support by Group

Group	Mean	N	Std. Deviation
TEL	3.41	27	0.89
IPT	4.46	28	0.64
Total	3.95	55	0.93

Table 12. *Table summarizing average support for Proposition A by group. The IPT group were more willing to sign the petition than the TEL group (p <.001).*

Town Hall Meeting. Although more participants in the IPT received and responded to the invite than those in the TEL group, 67% (13 out of 18) of participants in the IPT group accepted the invite as compared to only 36% (5 out of 15) in the TEL group. Moreover, more participants in the TEL group did not respond at all (36%) compared to those in the experimental group (11%). And finally, % of participants in the TEL group declined the invite, as compared to 6% in the IPT group.
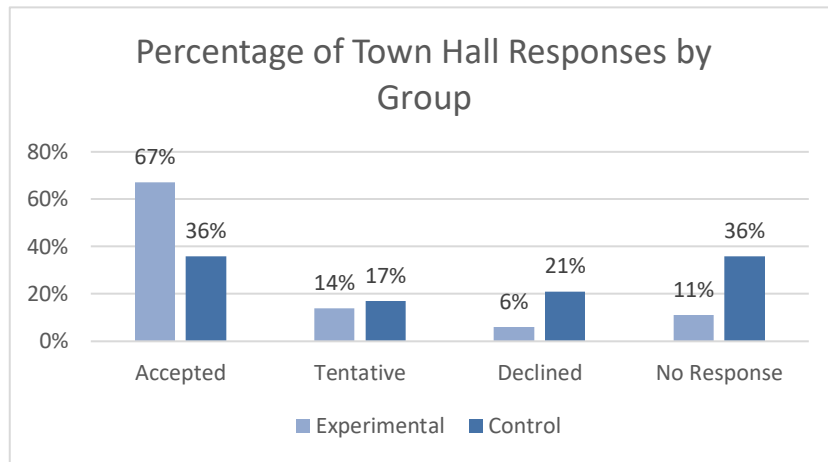

Figure 2. *Percentages of four responses to the town hall invite were divided by experimental group. IPT is the experimental group and TEL is the control group.*

4   DISCUSSION

This study explored the efficacy of two versions of the Hear to Listen course as an intervention to improve mental health literacy, empathy, and advocacy. Outcome variables included changes in positive and negative attitudes towards mental illness, positive and negative empathic responses, cognitive empathy, willingness to engage with a person who has a mental health condition. Additionally, in order to test for real-life transference, we included behavioral measures with willingness to sign a policy proposition, a town hall invite, and two letter writing tasks.

Results confirmed our hypotheses that the IPT group performed stronger on these outcome and behavioral measures than the TEL even though both groups showed improvements that were sustained over the course of the 6-week longitudinal experiment.

Cognitive empathy. The IPT group not only had higher cognitive empathy on the IOS post-test than the TEL, but they also had a greater increase in their scores from pretest-posttest. Additionally, those in the IPT group who were lower in dispositional perspective-taking abilities have a more significant increase in their skills than those in the TEL group who struggle with perspective-taking. This indicates that the IPT exercise might be particularly useful to increase cognitive empathy for those who struggle with perspective-taking. Additionally, participants who were higher in trait personal distress in the IPT group had a greater change in cognitive empathy after the course than those high in personal distress in the TEL group. This indicates that the IPT intervention might be useful for individuals who struggle feeling personally distressed while empathizing.

Empathic Responses. Both groups had increases in positive empathy (empathic concern) and decreases in negative empathy (empathic distress) through the course of the intervention. However, compared to the TEL group, the IPT group had more significant increases in feeling softhearted and touched, as well as significant decreases in feeling distressed and troubled. This indicates that the perspective-taking task contributed towards helping participants to regulate their personal distress after the second conversation with Isaac and to also feel more empathic concern for him. As previous work has linked emotion regulation with higher empathic concern, this finding is consistent with those results (Lebowitz & Dovidio, 2015).

Attitude Change. Both groups decreased negative biases and increased positive attitudes towards mental illness through the course of the experiment. However, the IPT group had more significant decreases in the extent to which they blame mentally ill persons for their condition. The IPT group also had greater increases than the TEL group in positive affirming attitudes about the success and empowerment of persons with a mental health condition.

Willingness to Engage. Willingness to interact in close proximity and on a regular basis with someone experiencing mental health issues was increased in both groups after the intervention. Again, the IPT had greater self-reported gains in this willingness than the TEL group, and these gains were sustained over time at a higher level than the TEL group.

Advocacy. The participants in the IPT group reported stronger agreement and likelihood than the TEL group to sign a petition that would increase governmental spending for mental health services in the workplace. The IPT group also used language exhibiting prosocial intent in their letter writing much more strongly than the TEL group, and used words like "support" and "services" more in their writing. The IPT group also used fewer "I" words and more "we" words than the TEL group, indicating a community orientation.

Volunteering free time. The IPT group were much more likely than the TEL group to respond and commit to attending a virtual meeting that would continue the mental health literacy training. This involved a personal sacrifice, as it occurred during non-working hours.

5   CONCLUSION

Thoughtful and well-designed eLearning can make a real difference, even for challenging and complex topics such as mental health awareness. We saw changes in attitudes and understanding in both groups of participants. These results indicate that there are advantages to using perspective-taking in online learning. Allowing participants to record their own video and then to see themselves from their conversation partner's point of view has clear advantages in cultivating long-term empathy. Both groups exhibited strong gains in empathy after the course, but the IPT group was much more willing to act and advocate on behalf of mentally ill persons and mental health literacy. This highlights the importance of interactive learning, and effectively isolates perspective-taking as one key design technique that can be utilized to build long-term empathy and advocacy in learners. Whereas the TEL group were more passive observers, the IPT group were more active learners. This difference led to the IPT group expressing much stronger calls to action in their letter writing, and much higher motivation to sign a political proposition and volunteer to learn more about mental health.

REFERENCES

Ahn, S. J. (Grace), Le, A. M. T., & Bailenson, J. (2013). The Effect of Embodied Experiences on Self-Other Merging, Attitude, and Helping Behavior. *Media Psychology*, *16*(1), 7–38. https://doi.org/10.1080/15213269.2012.755877

Banakou, D., Hanumanthu, P. D., & Slater, M. (2016). Virtual Embodiment of White People in a Black Virtual Body Leads to a Sustained Reduction in Their Implicit Racial Bias. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00601

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective Taking: Imagining How Another Feels Versus Imaging How You Would Feel. *Personality and Social Psychology Bulletin*, *23*(7), 751–758. https://doi.org/10.1177/0146167297237008

Corrigan, P. W., Powell, K. J., & Michaels, P. J. (2014). Brief battery for measurement of stigmatizing versus affirming attitudes about mental illness. *Psychiatry Research*, *215*(2), 466–470. https://doi.org/10.1016/j.psychres.2013.12.006

Daoudi, I., Chebil, R., Tranvouez, E., Lejouad Chaari, W., & Espinasse, B. (2021). Improving Learners' Assessment and Evaluation in Crisis Management Serious Games: An Emotion-based Educational Data Mining Approach. *Entertainment Computing*, *38*, 100428. https://doi.org/10.1016/j.entcom.2021.100428

Eckert, D., & Mower, A. (2020). The effectiveness of virtual reality soft skills training in the enterprise: A study. *PricewaterhouseCoopers*. https://www.pwc.com/us/en/tech-effect/emerging-tech/virtual-reality-study.html

Evans-Lacko, S., Rose, D., Little, K., Flach, C., Rhydderch, D., Henderson, C., & Thornicroft, G. (2011). Development and psychometric properties of the Reported and Intended Behaviour Scale (RIBS): A stigma-related behaviour measure. *Epidemiology and Psychiatric Sciences*, *20*(3), 263–271. https://doi.org/10.1017/S2045796011000308

Fernando, A. T., & Consedine, N. S. (2014). Development and initial psychometric properties of the Barriers to Physician Compassion questionnaire. *Postgraduate Medical Journal*, *90*(1065), 388–395. https://doi.org/10.1136/postgradmedj-2013-132127

Foster, A., Chaudhary, N., Kim, T., Waller, J. L., Wong, J., Borish, M., Cordar, A., Lok, B., & Buckley, P. F. (2016). Using Virtual Patients to Teach Empathy: A Randomized Controlled Study to Enhance Medical Students' Empathic Communication. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *11*(3), 181–189. https://doi.org/10.1097/SIH.0000000000000142

Goldfarb, E., & Gorrindo, T. (2013). Simulation in psychiatry. In A. Levine, S. DeMaria, A. D. Schwartz, & A. J. Sim (Eds.), *The Comprehensive Textbook of Healthcare Simulation* (pp. 522–523). Springer.

Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLOS ONE*, *13*(10), e0204494. https://doi.org/10.1371/journal.pone.0204494

Hirsh, A. T., Miller, M. M., Hollingshead, N. A., Anastas, T., Carnell, S. T., Lok, B. C., Chu, C., Zhang, Y., Robinson, M. E., Kroenke, K., & Ashburn-Nardo, L. (2019). *A randomized controlled trial testing a virtual perspective-taking intervention to reduce race and socioeconomic status disparities in pain care*. *160*(10), 12.

Jorm, A. F., Korten, A. E., Jacomb, P. A., Christensen, H., Rodgers, B., & Pollitt, P. (1997). "Mental health literacy": A survey of the public's ability to recognise mental disorders and their beliefs about the effectiveness of treatment. *Medical Journal of Australia*, *166*(4), 182–186. https://doi.org/10.5694/j.1326-5377.1997.tb140071.x

Kitchener, B. A., & Jorm, A. F. (2008). Mental Health First Aid: An international programme for early intervention. *Early Intervention in Psychiatry*, *2*(1), 55–61. https://doi.org/10.1111/j.1751-7893.2007.00056.x

Lebowitz, M. S., & Dovidio, J. F. (2015). Implications of emotion regulation strategies for empathic concern, social attitudes, and helping behavior. *Emotion*, *15*(2), 187–194. https://doi.org/10.1037/a0038820

Martingano, A. J., Hererra, F., & Konrath, S. (2021). Virtual Reality Improves Emotional but Not Cognitive Empathy: A Meta-Analysis. *Technology, Mind, and Behavior*, *2*(1). https://doi.org/10.1037/tmb0000034

McManus, S., Bebbington, P., Jenkins, R., Brugha, T., NHS Digital, & UK Statistics Authority. (2016). *Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014 : a survey carried out for NHS Digital by NatCen Social Research and the Department of Health Sciences, University of Leicester*.

Moll, S., Patten, S. B., Stuart, H., Kirsh, B., & MacDermid, J. C. (2015). Beyond silence: Protocol for a randomized parallel-group trial comparing two approaches to workplace mental health education for healthcare employees. *BMC Medical Education*, *15*(1), 78. https://doi.org/10.1186/s12909-015-0363-9

Morgan, A. J., Ross, A., & Reavley, N. J. (2018). Systematic review and meta-analysis of Mental Health First Aid training: Effects on knowledge, stigma, and helping behaviour. *PLOS ONE*, *13*(5), e0197102. https://doi.org/10.1371/journal.pone.0197102

Narayanasamy, M. J., Thomson, L., Coole, C., Nouri, F., & Drummond, A. (2020). Investigating the barriers and facilitators to implementing mental health first aid in the workplace: A qualitative study. *The Journal of Mental Health Training, Education and Practice*, *16*(2), 164–178. https://doi.org/10.1108/JMHTEP-12-2019-0064

Ojanperä, E. (2021). *Learning experiences and effects of online soft skills trainings*. 93.

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–787. https://doi.org/10.1016/j.concog.2013.04.016

Petkova, V. I., & Ehrsson, H. H. (2008). If I Were You: Perceptual Illusion of Body Swapping. *PLOS ONE*, *3*(12), e3832. https://doi.org/10.1371/journal.pone.0003832

Schumann, K., Zaki, J., & Dweck, C. S. (2014). Addressing the empathy deficit: Beliefs about the malleability of empathy predict effortful responses when empathy is challenging. *Journal of Personality and Social Psychology*, *107*(3), 475–493. https://doi.org/10.1037/a0036738

Seinfeld, S., Arroyo-Palacios, J., Iruretagoyena, G., Hortensius, R., Zapata, L. E., Borland, D., de Gelder, B., Slater, M., & Sanchez-Vives, M. V. (2018). Offenders become the victim in virtual reality: Impact of changing perspective in domestic violence. *Scientific Reports*, *8*(1), 2692. https://doi.org/10.1038/s41598-018-19987-7

Stevenson, D., & Farmer, P. (2017). *Thriving at Work: A review of mental health and employers*. https://www.gov.uk/government/publications/thriving-at-work-a-review-of-mental-health-and-employers

Zigman, D., Young, M., & Chalk, C. (2013). Using Simulation to Train Junior Psychiatry Residents to Work With Agitated Patients: A Pilot Study. *Academic Psychiatry*, *37*(1), 38. https://doi.org/10.1176/appi.ap.11070129