The Effects of a Live Versus an Automated Interviewer on the Emotional Engagement of Interviewees and their Perceived Empathy of the Interviewer

Anna-Karin Noromies, 34987
Master's Thesis in Psychology
Supervisor: Thomas Nyman
Faculty of Arts, Psychology and Theology
Åbo Akademi University
2019

Table of Contents

ÅBO AKADEMI UNIVERSITY – FAKULTY OF ARTS, PSYCHOLOGY AND THEOLOGY

Abstract for Master's Thesis

Subject: Psychology
Author: Anna-Karin Noromies
Title: The Effects of a Live Versus an Automated Interviewer on the Emotional Engagement of Interviewees and their Perceived Empathy of the Interviewer
Supervisor: Thomas Nyman
Abstract: In clinical settings, social presence, empathy and emotional engagement are related to positive outcomes. However, in online settings little is known about how these factors are perceived or if there are differences between a live interaction (e.g., a skype call) versus a video-based interaction (e.g., pre-recorded videos instructed by a test instructor). Here, in a short online face-to-face interview via a computer, we experimentally investigated the effect of automation of the interviewer's presence on the interviewee's emotional engagement and perceived empathy and presence of the interviewer. We defined degrees of automation (i.e., less rich social cues) by type of presence and the flexibility of the interviewer's response options, resulting in three experimental conditions: 1) a semi-scripted interview (live), 2) a scripted interview (live-scripted), and 3) a scripted interview where pre-recorded video clips substituted the interviewer (pre-recorded-video). A total of 75 university students participated and we found that in the live condition *(n* = 25) and the live-scripted condition (*n* = 25), participants expressed more joy ($F(2,75) = 5.49$, $p < .01$ partial $\eta^2 = .132$) and rated the interviewer as more empathetic ($V = .510$; $F(20,128) = 2.191$, $p = .005$; (partial) $\eta^2 = .255$) compared to the pre-recorded video condition (*n* = 25), whereas the live and the live-scripted condition did not differ in expressions of joy or perceived empathy. Perceiving the interviewer in the pre-recorded video as less realistic may have resulted in this outcome. In contrast, self-reported affective engagement did not differ between conditions. We found positive correlations between perceptions of empathy and activation of affect ($r = .32$; $p = .006$) and joy ($r = .29$; $p = .01$), suggesting that more perceived empathy resulted in a more positive experience. Together, our results indicate that a video-based interaction is perceived and experienced differently than a live interaction.
Keywords: automation, emotional engagement, empathy, social presence
Date: 13.11.2019 | Page count: 44

ÅBO AKADEMI – FAKULTETEN FÖR HUMANIORA, PSYKOLOGI OCH TEOLOGI
Abstrakt för avhandling pro gradu

Ämne: Psykologi
Författare: Anna-Karin Noromies
Arbetets titel: Effekterna av en live- versus automatiserad intervjuare på de intervjuades emotionella engagemang och upplevelse av empati från intervjuaren
Handledare: Thomas Nyman
Abstrakt: Empati, social närvaro och emotionellt engagemang har visats sig vara kopplad till positiva utfall i kliniska sammanhang. Hur dessa faktorer upplevs online eller om det finns skillnader mellan en live interaktion (t.ex., ett skypesamtal) och en videobaserad interaktion (t.ex., inspelade videoklipp som styrs av en testledare) är outforskat. I föreliggande experimentella studie, där 75 universitetsstuderande deltog i en datorbaserad intervju ansikte-mot-ansikte online, undersökte vi effekten av automatiserad närvaro av intervjuaren på den intervjuades emotionella engagemang och upplevelse av närvaro och empati från intervjuaren. Vi definierade automatiseringsgrad utifrån typ av närvaro och flexibiliteten i intervjuarens svarsalternativ. Studien innefattade följande tre betingelser: 1) en semiostrukturerad intervju (live), 2) en skriptad intervju (skriptad live), och 3) en skriptad intervju där intervjuaren ersatts med inspelade videoklipp (inspelad video). Vi fann att deltagarna i livebetingelsen (n = 25) och i den skriptade livebetingelsen (n = 25) oftare visade glädje (F(2,75) = 5.49, p < .01; (partiell) η^2 = .132) och bedömde intervjuaren som mer empatisk (V = .510; F(20,128) = 2.191, p = .005; (partiell) η^2 = .255) jämfört med deltagarna i den inspelade videobetingelsen (n = 25). Däremot fanns det ingen skillnad i uttryckt glädje eller upplevd empati mellan deltagarna i livebetingelsen och den skriptade livebetingelsen. Resultatet kan bero på att den inspelade videointervjun upplevdes som mindre realistisk än liveintervjuerna. Vi fann heller inga skillnader mellan betingelserna gällande deltagarnas självskattade emotionella engagemang. Vi fann att upplevd empati korrelerade positivt med affektens aktiveringsgrad (r = .32; p = .006) och uttryckt glädje (r = .29; p = .01), vilket kan indikera att intervjun gav en mer positiv upplevelse om intervjuaren upplevdes som empatisk. Sammanfattningsvis tyder resultaten på att en videobaserad interaktion uppfattas och upplevs annorlunda än en liveinteraktion.
Nyckelord: automatisering, emotionellt engagemang, empati, social närvaro
Datum: 13.11.2019 | Sidoantal: 44

Acknowledgements

In Åbo, November 2019,

Firstly, I would like to express my deepest appreciation to my supervisor Thomas Nyman. Your engagement and guidance in the thesis process have been valuable for my own engagement throughout. Thank you for all help with technical problems and other challenges I have come across in the experiment. Thank you for continuously being accessible, either in person, through skype, or via e-mail correspondences, and lowering the threshold to ask for help, something I often find challenging. Your regular check-ins kept me going when I was struggling the most, and your constructive feedback has made me aware of what I have learned and what I have yet to learn, without focusing too much on the latter. Your support has facilitated my confidence and helped me gain perspective on the process, that in the end it is just a text.

Furthermore, I would like to acknowledge the people participating in both the HumeLab seminars and Matti Laine & Mira Karrasch' seminars. Thank you for commenting on my work, and giving me the reader's perspective. I also want to thank Jan Antfolk for the help with the statistical analyses. Moreover, I would like to thank Sarah Haga for proofreading my Swedish summary. Last, but not least I would like to thank the people dearest in my life. Your support and presence in the process have been extremely valuable to me.

The Effects of a Live Versus an Automated Interviewer on the Emotional Engagement of Interviewees and their Perceived Empathy of the Interviewer

The global prevalence of mental disorders is high and continues to grow. The current estimate of individuals that suffer from the most common mental disorders, depression and anxiety, is 322 and 264 million respectively (World Health Organization, 2017). Less than half of those suffering from a disorder receive help due to different barriers to treatment, such as travel or treatment expenses, lack of resources, and stigma (Harvey & Gumport, 2015; World Health Organization, 2017). Therefore, finding ways to reach out to those in need of treatment is becoming increasingly important. Internet-based technology offers ways to overcome some of the barriers that can hinder the possibility of receiving face-to-face treatment, prompting research on internet-based mental health interventions (Barak & Grohol, 2011). To date, one of the most widely studied online methods is internet-based cognitive behavioural therapy (iCBT), which has been found to be as effective as face-to-face interventions for a variety of different psychiatric conditions (such as phobias and social anxiety), demonstrating both short-term and long-term effects (Andersson, 2018). This approach is based on cognitive behavioural therapy, an evidence-based treatment method that focuses on improving the way an individual feels and helps them overcome psychological problems by challenging unhelpful thoughts and behaviours (INSERM, 2004; Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012).

Research on internet-based interventions indicates that the effectiveness and adherence to internet-based interventions can be improved by human support (Andersson, 2009). Nevertheless, studies on unguided or automated internet-based interventions delivered by a computer or mobile phone with no human input or presence, have also been found to be effective (e.g., Griffiths, Farrer & Christensen, 2010; Titov et al., 2008; Titov et al., 2009a). It has even been found that such methods can have long-term effects (Vedaa, 2019). However, since to date, adherence to unguided interventions is low (Christensen, Griffiths, & Farrer, 2009; Richards & Richardson, 2012; Lillevoll et al., 2014), attempts have been made to improve automated interventions by including a simulated human presence (i.e., a virtual representation of a psychologist) and this have been proven to foster a therapeutic relationship with the program (Pinto et al., 2013, 2016; Rehm et al., 2016). For example, Pinto and colleagues (2013, 2016) found that participants experienced a sense of rapport and presence with a simulated avatar health-care professional and the group who interacted with the avatar had significantly fewer

depression symptoms compared to the attentional control group (receiving a computer-based health education). Pinto and colleagues (2016) speculated that this could be related to the interaction with the avatar. In traditional face-to-face therapies, the therapeutic relationship formed with the therapists predicts a positive treatment outcome (Norcross & Lambert, 2018) and it is likely that this is also an important factor in internet-based interventions. However, research on online methods, such as iCBT, indicates that online approaches may not always require live human support to be effective or that simulated human presence can be as effective as a traditional method. This highlights the need to further understand the role that human presence plays in an online treatment process and what the minimum amount and form of presence that is needed in order for a treatment to be effective. Indeed, while guided online interventions give better outcomes, and are favoured by clients (Holst et al., 2017), they are less scalable, more expensive and more difficult to implement compared to unguided interventions (Leykin, Muñoz, Contreras, & Latham, 2014). Automated programs are an attractive alternative to traditional methods since they could be a cost-effective alternative offered to individuals that may otherwise be placed on a waiting list (Cujipers et al., 2011). Nevertheless, although internet-based interventions show promise as alternative methods for treating less complex psychological problems, they should perhaps not be viewed as a replacement for traditional treatment methods. This is because online methods are not necessarily a good fit for everyone and may not be suitable for more severe psychological problems (Andersson & Titov, 2014).

Based on the need to better understand how a simulated or automated human presence is perceived by individuals in an online context, we decided to investigate how the automation of a human presence in a 10-15-minute online interview affects factors that are central to the therapeutic alliance; that is, the emotional engagement of interviewees and their perceived empathy of the interviewer. To achieve this aim, we created three conditions ranging from a normal online interview (live), to a more rigid interview (live-scripted), to a fully automated condition (pre-recorded video). In each interview condition we measured the reactions of the interviewees and their perceptions of the interviewer.

The Therapeutic Alliance, Emotional Engagement and Perceived Empathy in Clinical Interactions.

It has been shown that a strong therapeutic alliance or relationship (i.e., the emotional bond and agreement on tasks and goals of the treatment between the client and the therapist) is

an important predictor of positive outcomes in traditional face-to-face psychotherapy and counselling (Norcross & Lambert, 2018). A central element of the therapeutic alliance is the emotional engagement of both the therapist and the client, and it has also been shown that the therapists' empathic ability is a predictor of a client's change (Elliot et al., 2018; Norcross & Lambert, 2018). Moreover, the perception of a therapist as empathetic contributes to the development of a positive therapeutic alliance (Nienhuis et al., 2018, Watson et al., 2014) and promotes therapeutic change by facilitating client initiative, social interaction, and engagement (Watson et al., 2014).

In studies on guided internet-based interventions, it has been found that the working alliance (i.e., the agreement on tasks and goals of the treatment) has been rated to be as positive and as stable as in face-to-face interactions (Knaevelsrud and Maercker, 2006; Preschl et al., 2011; Wagner et al., 2012). However, in internet-based interventions the working alliance can rarely be considered to be a predictor of a positive outcome (Berger, 2017; Andersson, 2018). For example, Knaevelsrud and Maercker (2006) found that the alliance in a guided iCBT program for posttraumatic stress reactions was not as clearly related to the outcome compared to face-to-face approaches, despite high ratings of a therapeutic alliance. It has also been found that a client can develop a relationship with a computer program (e.g., Bickmore, Gruber, & Picard, 2005; Bickmore, Caruso, Clough-Gorr, & Heeren, 2005; Ormrod et al, 2010). This relationship is possible since people tend to treat computers as social beings, a tendency that is known as the Computers as Social Actors (CASA) paradigm, which implies that an individual reacts to the social cues provided by a computer, an avatar, or an algorithm (e.g., Reeves & Nass, 1996, Nass & Moon, 2000). While people elicit social responses to computers, little is known whether computer interactions are perceived as empathetic and emotionally engaging. As computer interactions are becoming more used in clinical interactions, knowledge about how they are perceived and experienced is needed due to their importance in clinical interactions.

Social Presence in Clinical Interactions.

An essential aspect of the human-computer interaction is the subjective experience of contact with the real or artificial other, also referred to as social presence by Lee (2004). Social presence has been shown to contribute to positive communication outcomes in mediated environments (Oh, Balieson & Welch, 2018), thus making it relevant to the design of computer programs in health care and technologies simulating clinical interactions (Lee, 2004). A

technological feature that influences the sense of social presence is the communication modality used (Oh, Bailenson and Welch, 2018). It has, for example, been found that text-based computer-mediated communication (CMC) evokes less social presence than richer forms of media (e.g., video, audio or avatar) (e.g., Bente et al., 2008). Contrasting studies have, however, shown that with longer interactions even media with richer social cues can evoke as much sense of social presence and contact that is comparable to face-to-face interaction, because individuals adapt to less rich social cues and take on other communication strategies such as direct questioning or self-disclosure (e.g., Ramirez et al., 2002; Antheunis et al., 2010, Walther, 1992).

Another important influential factor on social presence is visual representation. Previous research has concluded that the extent to which a visual representation acts the way a human would (i.e., behavioural realism), has positive effects on perceived social presence (e.g., Bente et al., 2008; Pan et al., 2008; Pütten et al., 2010). For example, Pütten and colleagues (2010) found that participants felt higher social presence when a computerized agent was nodding its head compared to an agent that was not. Similar effects have also been found for maintenance of mutual eye contact (Bente et al., 2008), and a virtual agent blushing strongly after making a mistake during a presentation (Pan et al., 2008). Researchers have also shown that the effect of behavioural realism is dependent on how the representation looks. For example, the more realistic a representation looks, the more realistic its behaviour needs to be in order for higher social presence to be evoked (Bailenson et al., 2005; Garau et al., 2003).

To date, there is no widely accepted, validated and generalized measure of presence across varied media or settings, due to the different existing conceptualizations of presence (Bailenson et al., 2005). Nevertheless, on social presence has often been operationalized either in terms of an individual's perceptions that another person is present or in terms of an individual's social response to the other (Bailenson et al., 2005). In the current study, we chose to measure social presence primarily by the extent to which the interviewer was perceived to be present and empathic, and secondarily as social responses, that is, the interviewee's self-rated experience of their own emotional engagement, and their emotional reactions measured through visual software analysis of their facial reactions.

The Current Study and Hypotheses

Although there is research that investigated how people respond to and evaluate interactions with computer representations (Oh, Bailenson & Welch, 2018), little is known about

how clinically relevant aspects of the interaction (i.e., empathy and emotional engagement) are perceived. The current study had two aims. First, in a 10-15-min online interview on wellbeing, we investigated if the perceptions and social responses of our participants differed when they interacted with an automated interviewer (i.e., pre-recorded video clips presented on a computer and directed by a test instructor) compared with when they interacted with a live semi-scripted online interviewer. Second, by creating a live condition where the interviewer had to strictly adhere to a script during the interview on wellbeing, we investigated if making the live online interview as similar (i.e., as rigid) as the pre-recorded video condition would have a similar effect on participants' responses as when the interviewer was automated. This resulted in three computer-based online interview conditions: 1) live, 2) live-scripted, 3) pre-recorded video.

In accordance with the CASA paradigm (Nass & Moon, 2000), we assumed that the interviewee would interact with the automated interviewer as if communicating with a live interviewer. Thus, we expected that the pre-recorded video condition would elicit similar reactions and emotional engagement in the participant as in the live condition, consistent with studies on social reactions to computers and avatars (e.g., Pütten et al., 2010, Qu et al., 2014). This was also based on the fact that the interviews were short and highly consistent. However, we also assumed that the live interaction would contain richer social cues compared to the automated (i.e., pre-recorded video) interaction and could thus evoke higher social presence. The rigidity of the interaction in the pre-recorded video condition might negatively affect the feeling of social presence, as the interaction is less spontaneous and natural. In sum, we expect the conditions to evoke similar responses, however, if the rigidity leads to less feeling of social presence, we should be able to detect differences between conditions. Second, we wanted to investigate whether self-rated emotional engagement, emotional reactions, and perceptions of empathy correlated positively. The reason for this was that previous research has shown that perceived empathy contributes to a positive alliance, and thus also the emotional engagement in the interaction (Nienhuis et al., 2018, Watson et al., 2014). In sum, we tested the following hypotheses:

Hypothesis 1: Increased automation of presence (i.e., less rich social cues) will decrease the emotional engagement, emotional reactions and perceived empathy, so that the live condition will evoke the highest ratings of perceived emotional engagement, emotional reactions, and perceptions of empathy, followed by the live-scripted condition and last the pre-recorded video

condition. Alternatively, there will be no differences in emotional engagement, emotional reactions, and perceptions of empathy between the conditions.

Hypothesis 2: Perceived emotional engagement, emotional reactions, and perceptions of empathy will be positively correlated.

Method

Participants

The participants were contacted through mailing lists at universities and vocational schools in Turku, through social media, and by recruitment at the university canteen. A sample of 75 Swedish-speaking students at the Åbo Akademi University participated in the study. Forty-nine participants were women and 25 were men. One participant did not provide any background information. The age of the participants ranged from 19 to 52 years with a mean age of 23.4 (*SD* = 4.09) years. The participants were pseudo-randomized into three separate experimental conditions. The live group (n = 25) had a mean age of 24.0 (*SD* = 6.17*),* the live-scripted group (n = 25) had a mean age of 23.2 (*SD* = 2.77), and the pre-recorded video group (n = 25) had a mean age of 22.8 (*SD* = 2.11). Participants received a lunch coupon as compensation for their participation.

The current study was approved by the Ethics committee at the Department of Psychology and Logopedics at Åbo Akademi University.

Design

The experiment setup was a between-subjects design with three conditions. The independent variable, automation of presence (i.e., richness of social cues), was operationalized as the flexibility of the interviewer's response options and form of presence (live vs. live-scripted vs. pre-recorded video). The dependent variables were participant ratings of the interviewer's social presence and empathic response, and their own emotional engagement. The three conditions were, (1) an online interview where test subjects saw a psychology student live through a computer screen and were interviewed verbally in accordance with a pre-defined list of lines and questions following a fixed order. In this condition, the interviewer (i.e., the psychology student) could freely validate or paraphrase non-verbally and verbally to the interviewee's responses before continuing to the next question. (2) An online interview where the test subject saw the psychology student live through a computer screen and was verbally interviewed in the same way as in condition 1, but here the interviewer was verbally limited to

six pre-defined answers (i.e., a script). (3) A video-based online interview with multimodal input in the form of video recordings of the interviewer. Here, also each pre-recorded line and question followed the same order as in condition 1 and 2, and the answers were pre-defined as in condition 2. The test instructor here was the interviewer who managed the video recordings.

Measures

In order to assess the emotional engagement of the participants, we measured self-reported emotional engagement and facial expressions. We assessed the participant's perceived emotional state using the self-report measure Swedish Core and Affect Scale (SCAS), administered before and after the interview. The SCAS measures the valence and activation of the current mood of the individual based on an affect grid (Västfjäll et al., 2002). The items are graded on a 9-point Likert-type scale ranging from –4 to +4, with the end points of the scale defined by adjective pairs (e. g., sad-glad, bored-interested, engaged-disengaged). We used the composite values of valence and activation. The SCAS measure shows adequate reliability and validity for composite scores of valence and activation (Västfjäll & Gärling, 2007). We measured participants' emotional reactions in the interview with automated facial expression analysis. Facial expression analysis provides a simple, low-time consuming, and less intrusive measurement of emotion (iMotions, 2016; Meiselman, 2016) and shows adequate sensitivity to the valence of emotional states (Russell, 1994, Mauss et al., 2005). This analysis was performed with the use of the AFFDEX Software Development Kit (SDK) 2.0. Using the built-in computer camera, the program software recognizes facial action units (AU) (i.e., facial movements) at each time point and classifies specific combinations of them as emotions, based on a facial coding scheme (Ekman & Friesen, 1976). A validation study of the Affdex software showed that in Study 1 the overall accuracy of valence was 73 % for prototypical facial expressions (picture) and 53% in Study 2 where more natural dynamic facial expressions (video) used. Cohen's cappa was 0.68 for Study 1 and 0.47 for Study 2 (Stöckl et al., 2018). The emotions take on values from 0 to +100. We calculated the mean value per all seven emotions measured in the AFFDEX SDK. The emotions measured were anger, sadness, disgust, joy, surprise, fear and contempt.

We measured the participants' perceptions of and reactions to the interviewer using the Consultation and Relational Empathy (CARE) measure (Mercer et al., 2004, Swedish translation). The CARE measure is considered to have high acceptability, face validity and internal construct validity (Ahlforn, Horwitz, & Osika, 2017) and is designed to evaluate the

extent to which the patient judges the consultant to be empathic in a one-on-one consultation. The items are graded on a six-point scale ranging from 1 (poor) to 5 (excellent), including a Not Applicable (NA) option (Kersten, White & Tennant, 2011).

We used the Patient Health Questionnaire (PHQ-9) (Kroenke, Spitzer & Williams, 2001, Swedish version) to investigate the intensity of symptoms of depression during the past two weeks, since depression has been shown to reduce emotional responses (Bylsma et al., 2008). We chose the PHQ-9 measure since it is suitable as a screening tool for depression in research settings and shows good validity and reliability (Kroenke, Spizter & Williams, 2001). The answers are scored on a four-point Likert-type scale ranging from 0 (not at all) to 3 (almost every day), with a total score ranging from 0 to 27. A total score of ≥ 10 was used as a cut-off score to for clinically significant depression symptoms (Kroenke, Spitzer & Williams, 2001). In addition, we also collected the following demographic data: gender, age, as well as previous experience of counselling or psychotherapy. Information of previous experience of counselling of psychotherapy was collected since experience of interaction with a mental health professional could form the participants' expectations of the interview and also affect the participant's perceptions of and reactions to the interviewer. The demographic data were obtained on a form labelled the background questionnaire.

Materials

We used an iMac (24-inch, 1920x1200, early 2009) and a HP Compac 8200 Elite MT PC to conduct the interviews in all three conditions. To conduct the live interviews, we used the website Doxy.me - © Doxy.me, LLC 2014. Doxy.me was chosen since it is considered to comply with HIPAA (the Health Insurance Portability and Accountability Act) and HITECH (the Health Information Technology for Economic and Clinical Health). HIPAA is a standard for privacy and security rules that protect patient data.

For the video interview, we filmed the interview questions and responses using a digital camera (Canon EOS 1300D; lens: Canon EF 50mm f/1.8). The video recordings captured the psychology student from shoulder-height up and were edited with OpenShot™ Video Editor. In total, the recorded material consisted of 27 video clips, which included four clips of the introductory part of the interview, 14 clips of the main interview questions, one clip that functioned as a bridge when changing the theme in the interview, one clip that was the concluding part of the interview, six clips representing the different interviewer responses, and

one clip that functioned as a dynamic background clip. To simulate the live scripted interview, we further edited the video-clips and then run them using Resolume Arena 5 so that the video clips could transition seamlessly. In the Resolume-Arena-5 software tool we created a composition of video clips organized in layers and in the desired order. The test instructor operated this simulation remotely.

Procedure

Before conducting the experiment, we piloted the experimental design in a small sample (n = 10) of students who were not included in the main study. We found no previously validated structured clinical interview for our experiment, so we devised a draft version and tested the interview to establish its comprehensibility and functionality, and the functionality of the pre-recorded video interview, and no alterations were made between the pilot and the main experiment. None of the participants reported any disturbing or irritating aspects of the pre-recorded video interview, however, for the live interviews some reported being disturbed by occasional desynchronization of audio and graphics, which was likely due to poor internet connection. We conducted the interviews in rooms at the Psychology Department at Åbo Akademi University. We asked the participants to take part in an approximately ten-to-fifteen-minute online interview concerning their health and everyday life. At that point they were given incomplete information regarding the subject of the study, and we did not reveal the specific hypothesis concerning the effect of automation of presence (i.e., less rich social cues) on emotional engagement and perceived empathy. This was necessary in order for us to test our hypothesis in a valid manner. The participants were assured anonymity and were informed that participation was voluntary and could be discontinued at any time. Upon arrival, a research assistant welcomed and gave each participant instructions. The setup was a double-blind as the assistant was unaware of the research question, the study design, and the group to which the participant had been assigned. The participants read and agreed to a standard consent form before starting the experiment.

Once seated in the test room, the participant first filled in a background information questionnaire. After this, they completed the Patient Health Questionnaire as well as Swedish Core and Affect Scale. Then, they repositioned themselves in front of the iMac computer at the adjoining table, put on headphones and waited for the interview to begin. We used the Doxy.me-platform and the AFFDEX software program on the iMac. Participants used sound-dampening

headphones to hinder hearing the interviewer talk, who was positioned in the adjacent test room. The research assistant instructed the participants to avoid changing sitting position and turning their head during the interview. The reason for this was to enable AFFDEX identify facial expressions and analyze them as accurately as possible. In every condition, the interview followed the same outline, thus allowing for comparisons between groups (see Appendix). The introductory part of the interview was to establish contact, explaining the aims of the interview, and restating that their participation could be discontinued at any time without explanation, and obtaining verbal affirmation to continue. The main part of the interview consisted of open questions (e.g., "Tell me about a typical day in your life", "How would you describe your overall health?", "Tell me a little bit about the most important persons in your life") relating to central themes discussed in an intake interview (Sommers-Flanagan & Shaw, 2017). In the live online interview, the interviewer was to a limited extent free to choose how to respond to the participant's answer. We operationalized the term "free" as spontaneously being able to validate or paraphrase what the participant has said (e.g., "You say that...", "It seems like you think that…", "If I understand you correctly you say that…"). The second and third conditions differed in the number of ways the interviewer could respond, as in these two conditions, the interviewer's answer followed a predesigned decision tree based on the response of the participant (see Figure 1).

At post-interview, the participant completed the SCAS questionnaire and the Consultation and Relational Empathy questionnaire. Immediately after completion, the research assistant thanked them for their participation and gave the study participants a debriefing statement. The debriefing form contained information of the real purpose of the study (however, the hypotheses were not disclosed), relevant background information pertaining the study, the researcher's contact information for follow-up questions and study results, and information about counselling services in Turku.

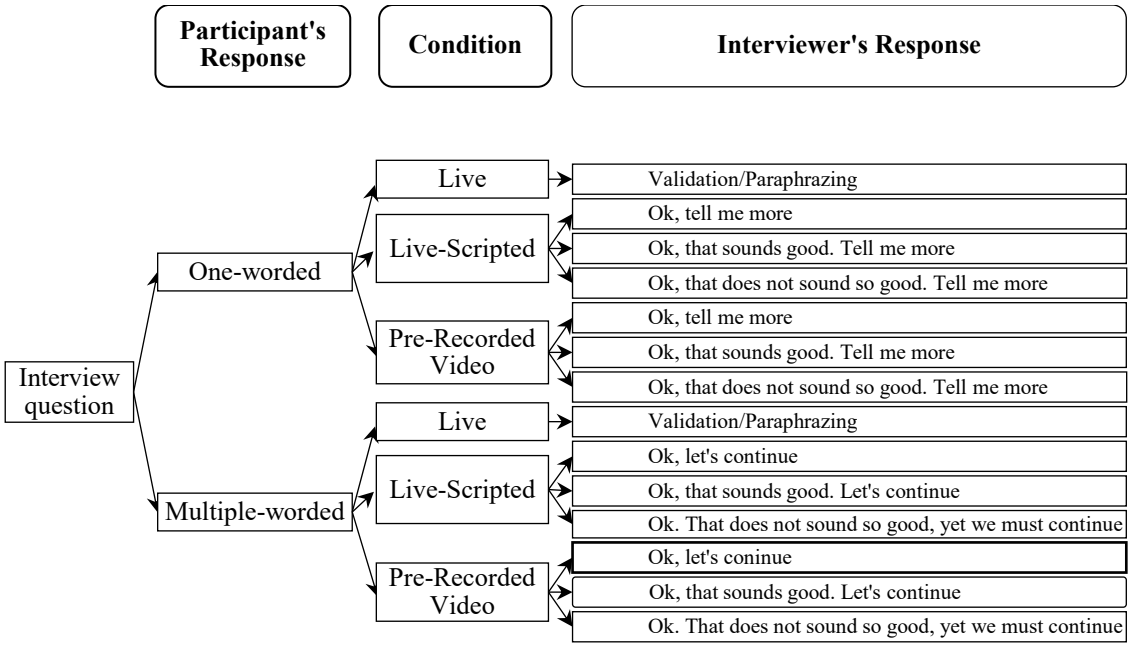

Figure 1. Schematic of the response decision tree. At the live interview, each response of the participant was followed by a validating or paraphrasing line. At the live-scripted and pre-recorded video interview, each response of the participant was followed by either a line trying to engage the participant or a line indicating that the interview continues on to the next question or stage of the interview.

Statistical Analyses

We used SPSS Statistics 24 for all our analyses. To test whether the participants were emotionally engaged in the pre-recorded video interview and if reported own engagement differed between groups before and after the interview, we performed a mixed-model ANOVA with time as the within-subject factor and group as the between-subject factor. We used the Valance Composite and Activation Composite of the SCAS measure as dependent variables. To then test if the total amount of expressed emotional reactions differed between groups and if there were differences between groups on each emotion variable, we conducted a MANOVA on all AFFDEX data, followed by one-way ANOVAs on each emotion separately. The reasons for retaining the outliers in the AFFDEX data was that the extreme values could represent an aspect of the inherent variability of the data (i.e., these outliers can be legitimate emotional reactions in an online setting). The dependent variables were the average of each of the seven measured emotions. To then investigate differences between groups concerning perceived empathy of the interviewer, we performed a MANOVA on the CARE measure. Lastly, to test if self-rated emotional engagement, emotional reactions, and perceptions of empathy of the interviewer were related, we conducted correlational analyses between the average score of the CARE measure,

the difference in Valence and Activation Composites pre- and post-interview, and the average of each expressed emotion.

Results

Data Preparation

Prior to the main analyses, we checked that the conditions were balanced and each condition had twenty-five subjects. We then investigated if there were any group differences based on the PHQ-9 scores, gender and previous experience. A Kruskal-Wallis H test revealed that group means of PHQ-9 scores did not differ significantly, $\chi^2(2) = 0.691$, $p = .708$, with a mean rank score of 40.80 for the live interview, 35.80 for the live-scripted interview, and 37.40 for the pre-recorded video interview. A Chi-Square test showed that the experience of psychotherapy or counselling did not differ significantly between groups, $\chi^2(2, N = 74) = 0.239$, $p = .888$. Likewise, gender was equally distributed across groups, $\chi^2(2, N = 74) = 2.663$, $p = .264$. Accordingly, these variables were excluded in subsequent analyses.

Next, we tested for missing data and found that 49 (9%) of the values from the AFFDEX dataset were missing. Analyses of missing data patterns showed that participants in the live-scripted group had five cases with missing values, the live group had two and the pre-recorded video group none along all variables. The data loss was a result of computer software failure. The AFFDEX software stopped in cases where the program was started before signing in to the website Doxy.me - © Doxy.me, LLC 2014. Since we used Doxy.me in only the live and live-scripted interviews, this software failure was only possible in these conditions, thus related to the group variable, indicating that the missing data is not missing completely at random, according to Rubin (1976) and colleagues' (Little & Rubin, 2015) missing data theory. In order to retain statistical power, we addressed the missing values for each variable by imputation with the mean of the sample, as multiple imputation can be applied to both data that are missing at random and not at random (Schafer & Graham, 2002). There CARE questionnaire had 128 (17%) Not Applicable (NA) responses. Most of the NA responses were checked for items 8 (Explaining things clearly), 9 (Helping you to take control), and 10 (Making a plan of action with you). We replaced these NA responses with the mean score of the rest of the valid responses according to the suggested scoring procedure (Mercer et. al, 2005). Although a fraction of the values was missing from the background variables (1%) and the SCAS questionnaire (2%), a missing rate up

to 5% non-significantly affects the quality of the analyses (Schafer, 1999). We handled the missingness by pairwise deletion.

The Effects of a Live Versus an Automated Interviewer on the Self-Rated Emotional Engagement of Interviewees

We investigated the effect of automation of presence (i.e., less rich social cues) on valence and activation of affect, controlling for pre-interview emotion ratings. The result of a one-way repeated measures ANOVA showed that the valence of affect was significantly more positive post-interview compared to pre-interview, $F(1,75) = 13.05$, $p < .001$. Contrary to our expectation, there was no significant difference in valence of affect between the groups, $F(2,75) = 0.63$, $p = .939$. Likewise, a separate one-way repeated ANOVA showed that the activation of affect was significantly higher post-interview compared to pre-interview, $F(1,75) = 35.81$, $p < .001$. Participants in the live interviews tended to be more emotionally activated compared to participants in the pre-recorded video interview, however, the difference in activation of affect did not differ significantly between groups, $F(2,75) = 2.88$, $p = .063$ (see Table 1 for means). The results suggest that the live, live-scripted, and pre-recorded video condition elicited similar emotional engagement, but that there was no difference in valence or activation of affect between groups.

Table 1

Results of Statistical Analyses for SCAS, AFFDEX, and CARE measure in Live, Live-Scripted and Pre-Recorded-Video Interview

	Live (l)	Live-Scripted (ls)	Pre-Recorded Video (v)				
Variable	*M (SD)*	*M (SD)*	*M (SD)*	*F*(2,75)	*p*	partial η^2	Tukey's HSD
SCAS							
PreVal	1.04 (1.57)	1.11 (1.13)	1.19 (1.49)				
PostVal	1.73 (1.59)	1.55 (1.17)	1.31 (1.67)				
PreAct	-0.04 (2.14)	0.82 (1.79)	1.24 (1.58)				
PostAct	1.23 (1.63)	1.97 (1.09)	1.77 (1.27)				
AFFDEX							
Anger	1.21 (3.05)	0.50 (0.57)	1.02 (1.42)	0.870	.417	.024	-
Contempt	4.98 (5.18)	5.13 (3.93)	5.81 (4.40)	0.236	.790	.007	-

Disgust	2.60 (3.85)	1.97 (2.67)	1.55 (1.93)	0.827	.441	.022	-
Fear	0.53 (0.56)	0.45 (0.41)	0.63 (0.91)	0.515	.600	.014	-
Joy**	15.1 (15.9)	17.2 (19.2)	4.24 (6.52)	5.49	<.01	.132	l, ls > v
Sadness	0.68 (1.40)	0.32 (0.27)	0.58 (0.80)	1.055	.353	.028	-
Surprise	6.57 (4.13)	7.58 (5.39)	5.15 (4.94)	1.58	.212	.042	-
CARE item							
1.Making you feel at ease***	4.28 (0.89)	4.00 (0.76)	3.08 (1.08)				
2.Letting you tell your story*	4.40 (0.65)	4.28 (0.74)	3.76 (1.16)				
3.Really listening	4.44 (0.92)	4.28 (0.92)	4.28 (0.84)				
4.Being interested in you as a whole person*	4.16 (1.14)	3.76 (1.01)	3.20 (1.32)				
5.Fully understanding your concerns**	3.92 (1.15)	3.52 (1.05)	2.76 (1.27)				
6.Showing care and compassion**	4.04 (1.17)	3.72 (0.98)	2.80 (1.22)				
7.Being positive***	4.28 (0.89)	4.32 (0.80)	3.28 (1.02)				
8. Explaining things clearly	4.16 (1.07)	4.24 (0.78)	4.00 (1.08)				
9. Helping you take control	3.88 (1.17)	3.76 (0.93)	3.24 (1.16)				
10. Making a plan of action with you*	4.08 (1.15)	3.80 (0.96)	3.28 (1.21)				
CARE average**	4.17 (0.89)	3.97 (0.70)	3.34 (0.91)				

Note. Live = semi-scripted face-to-face interview via a computer; Live-Scripted = scripted face-to-face interview via a computer; Pre-Recorded Video = pre-recorded video interview via a computer; PreVal = valence of self-rated emotion pre-interview; PreAct = activation of self-rated emotion pre-interview, PostVal = valence of self-rated emotion post-interview, PostAct = activation of self-rated emotion post-interview; Anger = amount of facial expressions of anger; Contempt = amount of facial expressions of contempt; Disgust = amount of facial expressions of disgust ; Fear = amount of facial expressions of fear; Joy = amount of facial expressions of joy; Sadness = amount of facial expressions of sadness; Surprise = amount of facial expressions of surprise; CARE = amount of perceived empathy. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

The Effects of a Live Versus an Automated Interviewer on the Emotional Reactions of Interviewees

We found no significant main effect of automation (i.e., less rich social cues) of presence on the amount of facial expressions, $F(14,134) = 1.153$, $p = .319$, $V = .215$, indicating that there was no difference in the amount of expressed emotion between the interview groups. Separate univariate ANOVAs revealed that groups only differed in the expression of joy, $F(2,75) = 5.49$, $p < .01$ partial $\eta^2 = .132$ (see Table 1). A post hoc Tukey's test comparing group means showed that the live group and the pre-recorded video group differed significantly. Also, the scripted-live group and the pre-recorded video group differed significantly. The live group

expressed most joy and the pre-recorded video group least joy. The live scripted group expressed more joy than the live group, but this difference was not significant (see Table 1). Our results suggest that total amount of emotional reactions did not differ between conditions, but that the live groups expressed more joy than the pre-recorded video group.

The Effects of a Live Versus an Automated Interviewer on the Interviewees' Perceived Empathy of the Interviewer

We aimed to discern whether automation of presence (i.e., less rich social cues) affected the interviewees' perceptions of the interviewer. Overall, people rated the live and the live-scripted interviewer as more empathic compared to the interviewer in the pre-recorded video. The MANOVA showed that there was a significant effect of automation on perceptions of empathy, $V = .510$; $F(20,128) = 2.191$, $p = .005$; (partial) $\eta^2 = .255$.

Separate univariate ANOVAs on the outcome variables, however, revealed a non-significant effect of automation on the perception of how well the interviewer listened, $F(2,75) = 1,532$, $p = .223$, explained things clearly, $F(2,75) = .384$, $p = .682$, or was helping one take control, $F(2,75) = 2.429$, $p = .095$. To determine which CARE variables best accounted for the difference, we followed up the MANOVA with a discriminant function analysis (DA) (Brown & Wicker, 2000). DA is predictive model that determines how many dimensions are needed to express the relationship between several independent variables (i.e., CARE items) and one categorical variable (i.e., group). Since results from separate ANOVAs indicated that group differences were determined by items 1,2,4,5,6,7, and 10, we included these variables in the discriminant analysis. The DA revealed two discriminant functions. We used Wilk's lambda, ranging from 0 to 1, to measure how much of the variance of the dependent variable was not explained by the discriminant functions. The first function explained 89.2% of the variance, canonical $R^2 = .37$, whereas the second explained only 10.8%, canonical $R^2 = .07$. In combination, the discriminant functions significantly differentiated the interview groups, $L = 0.594$, $\chi^2(20) = 35.892$, $p = .001$, but removing the first function indicated that the second did not significantly differentiate the interview groups, $L = .935$, $\chi^2(9) = 4.649$, $p = .590$. Therefore, the group differences shown by the MANOVA can be explained in terms of one underlying dimension (see Table 2).

Table 2

Summary of interpretive measures for discriminant analysis.

Independent variable	Unstandardized function		Standardized function		Discriminant loading (rank)	Univariate F ratio
	1	2	1	2		
CARE measure item						
1.Making you feel at ease	0.878	0.445	0.807	0.410	0.736 (1)	11.653***
2.Letting you tell your story	0.701	-0.418	0.617	-0.363	0.435 (5)	4.080*
4.Being interested in you as a whole person	-0.418	0.264	-0.487	0.308	0.417 (7)	4.275*
5.Fully understanding your concerns	0.084	0.692	0.098	0.801	0.530 (4)	6.467**
6.Showing care and compassion	0.262	0.195	0.296	0.221	0.610 (3)	8.108**
7.Being positive	0.746	-1.450	0.678	-1.319	0.713 (2)	10.500***
10. Making a plan of action with you	-0.888	0.257	-0.986	0.286	0.380 (6)	3.338*
Group centroid live					0.539	
Group centroid live scripted					0.509	
Group centroid pre-recorded video					-1.048	
Wilk's lambda (function 1 through 2)					0.594**	
(Canonical correlation)^2					0.367	

* $p < .05$; ** $p < .01$; *** $p < .001$.

Item 1 ($r = .74$) on the CARE measure had the greatest contribution to the first function. Most interesting is that on the first function, item 1, 2, 5, 6, 7 had positive weights, whereas item 4,10 had negative weights. This suggests that the group differences can best be explained by items 1, 2, 6 and 7 compared to the other items. The variate centroids for each group shows that variate 1 discriminates the pre-recorded video group from the live and live scripted combined because the pre-recorded video group have negative values (r = -1.048) whereas the live group (r = .539) and the live scripted group (r = .509) have positive values on the first function. As the positive weights suggest, the live group and the live scripted group were, however, not discriminated by the first variate.

The usefulness of the discriminant function was evaluated from accuracy of group

classification and cross-validation. Table 3 presents the classification results. It shows that the total accurate classification proportion of the discriminant function obtained was 61.33%. Moreover, the hit rate of discriminant analysis was higher than the base rate (61.33% > 33.33%), which means that the discriminant function provided correctly classified cases beyond chance classification. To determine the stability of the discriminant analysis statistics, we first performed a discriminant analysis on a cross-validated sample. The results show that the total hit rate was lower for the derived sample compared to the original sample (52.00% < 61.33%), indicating that the predictive accuracy is unstable.

Table 3

Classification Results for Original and Cross-Validated Sample

			Predicted Group Membership			
		Condition	Live	Live-Scripted	Pre-Recorded Video	Total
Original	Count	Live	15	6	4	25
		Live-Scripted	7	13	5	25
		Pre-Recorded Video	3	4	18	25
	%	Live	60.0	24.0	16.0	100
		Live-Scripted	28.0	52.0	20.0	100
		Pre-Recorded Video	12.0	16.0	72.0	100
Cross-validated	Count	Live	13	8	4	25
		Live-Scripted	10	10	5	25
		Pre-Recorded Video	5	4	16	25
	%	Live	52.0	32.0	16.0	100
		Live-Scripted	40.0	40.0	20.0	100
		Pre-Recorded Video	20.0	16.0	64.0	100

Note. 61.3 % of original grouped cases correctly classified. In the cross-validated sample, 52.0 % of cases were correctly classified.

Second, by comparing the hit rate to the chance rate we determined the significance of the discriminant function after cross-validation. The equation for this is

$$z = (Np_a - Np_c)/[Np_c(1 - p_c)], \quad (1)$$

where N is the total sample size, p_a the percentage of correctly classified cases using the discriminant analysis, and p_c the percentage of correctly classified cases with base rate.

Applying equation 1, $z = (75(.52)-75(.33)) / [75(0.33)\ (1-.033)] = 0.86$, which is lower than the critical value of 1.65 for an alpha level of .05. That is, the classification accuracy exceeds the expected on the basis of chance after cross-validation, however not at the .05 level of significance.

In sum, the different interview groups can be distinguished by several aspects of the interaction. In particular, ratings on how the interviewer was at making the participants feel at ease (i.e., introducing herself, explaining her position, being friendly, respectful, not cold or abrupt), and being positive (i.e., having a positive approach and a positive attitude; being honest but not negative about one's problems) discriminated the pre-recorded video group from the live and live scripted group, however not significantly. Our results suggest that the interviewer in the live and live-scripted was perceived as more positive and better at making the participants feel at ease compared to the interviewer in the pre-recorded video, but that the live and live-scripted group did not differ in these perceptions.

The Relationship between Emotional Engagement, Emotional Reactions and Perceived Empathy

Lastly, we tested if there was a positive correlation between perceived empathy, emotional engagement and emotional reactions (see Table 3). A correlational analysis showed that there was a significant positive correlation between perceived empathic response and positive change in activation of affect, $r = .32$; $p = .006$. We found a significant positive correlation between perceptions of empathy and expression of joy, $r = .29$; $p = .01$. The value of $R^2 = .10$, indicates that 10% of the variation in difference in activation of affect can be attributed to the variation in the empathy ratings, and the value of $R^2 = .08$, indicates that 8% of the variation in expressions of joy can be attributed to the variation in empathy ratings. While causality cannot be inferred based on these analyses, the results suggest that participants may have had a more positive experience (i.e., greater positive change in activation of affect and expressed more joy) when they perceived the interviewer as empathic. We also found a significant negative correlation between valence of affect and expressed fear, $r = -.25$; $p = .037$. This correlation could be a result of AFFDEX's poor ability to correctly identify fear in only 1 % of the cases (Stöckl et al., 2018), implicating that some of the other emotions (e.g., sadness) could have been mistaken for fear. The correlation analysis shows a strong correlation between fear and sadness, $r = .56$; $p < .001$. Positive intercorrelations between emotions on the AFFDEX

measures (see Table 3) were also found which suggests that participants expressing one emotion likely expressed also other emotions, that is, they were more emotionally engaged in the interview. One significant negative correlation was found between joy and contempt, $r = -.25$; $p = .035$, which is not that surprising since joy is a positively valanced emotion and contempt a negatively valanced emotion.

Table 3.

Summary of Intercorrelations, Means, and Standard Deviations for Scores on the SCAS, AFFDEX, and CARE measure

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1.DeltaVal	0.42	1.00	—									
2.DeltaAct	0.98	1.42	.34** [.09, .57]	—								
3.Anger	0.91	1.97	.03 [-.27, .18]	.04 [-.15, .35]	—							
4.Contempt	5.31	4.49	.11 [-.12, .38]	.05 [-.14, .25]	.12 [-.002, .41]	—						
5.Disgust	2.04	2.92	-.12 [-.40, .13]	-.09 [-.22, .10]	.20 [-.03, .56]	.24* [.001, .47]	—					
6.Fear	0.54	0.66	-.25* [-.42, -.03]	.18 [-.16, .46]	.56** [.32, .82]	.09 [-.19, .28]	.31** [.16, .54]	—				
7.Joy	12.20	15.76	.03 [-.24, .31]	-.18 [-.29, .06]	-.07 [-.21, .13]	-.25* [-.43, -.05]	.12 [-.23, .48]	-.07 [-.28, .20]	—			
8.Sadness	0.53	0.94	-.01 [-.31, .17]	.08 [-.12, .39]	.94** [.73, .98]	.21 [.05, .50]	.21 [-.04, .54]	.56** [.24, .77]	-.15 [-.29, -.04]	—		
9.Surprise	6.43	4.89	-.18 [-.44, .12]	-.11 [.005, .10]	-.08 [-.20, .08]	-.08 [.09, -.25]	.31** [-.002, .62]	.34** [.18, .57]	.35** [.09, .57]	-.10 [-.22, 05]	—	
10.CARE	3.82	0.90	.17 [-.16, .37]	.32** [.08, .50]	.12 [-.06, .24]	-.02 [-.22, .16]	-.01 [-.26, .24]	.10 [-.09, .30]	.29* [.10, .45]	.09 [-.11, .23]	.15 [-.10, .37]	—

Note. *$p < .05$, **$p < .001$. BCa bootstrap 95% Cls reported in brackets.

Discussion

In our experiment we set out to test how the automation of social presence (i.e., less rich social cues) would affect the interviewee's experience of the social presence, and perceived empathy of the interviewer, and the interviewee's emotional engagement during an online interview on wellbeing. We did this by manipulating the flexibility of an online interviewer's response options and the form of presence (i.e., live vs. live-scripted vs. pre-recorded). Our first goal was to investigate if the reactions and perceptions of our participants differed when they interacted with an automated interviewer (i.e., pre-recorded video clips) compared with a live online interviewer. Our second goal was to investigate if creating a condition where the live interviewer was forced to adhere to a script (i.e., making the live online interview as rigid as a pre-recorded video), would have similar effect on the reactions and perceptions of our participants as in the pre-recorded video condition. Based on the literature on human-computer interactions, that richer social cues increase social presence (Oh, Bailenson and Welch, 2018), we anticipated that increased automation of presence (i.e., less rich social cues) would decrease emotional engagement, reactions and perceived empathy, so that perceived emotional engagement, emotional reactions, and perceived empathy would be highest in the live condition, followed by the live-scripted condition and last the pre-recorded video condition. However, we also anticipated that the pre-recorded video condition would elicit similar reactions and emotional engagement in the participant as in the live condition, based on studies on social reactions to computers and avatars (e.g., Pütten et al., 2010, Qu et al., 2014) and that the interviews were short and highly consistent. Therefore, our alternative hypothesis was that we might not detect differences between conditions. We also wanted to analyse the relationship between self-rated emotional engagement, emotional reactions, and perceptions of empathy. Here, we anticipated that there would be a positive correlation between self-rated emotional engagement, emotional reactions, and perceptions of empathy, since perceived empathy has shown to contribute to a positive alliance, and thus also the emotional engagement in the interaction (Nienhuis et al., 2018, Watson et al., 2014).

The Effects of a Live Versus an Automated Interviewer on the Self-Rated Emotional Engagement of the Interviewees

The outcome of our analysis of emotional engagement indicated that automation had no significant effect on either activation or valence of affect. The average score of activation was

higher in the live conditions (live $M = 1.19$, live-scripted $M = 1.15$) compared to the pre-recorded video condition ($M = 0.53$) on a -4 to +4 scale. Although the differences were not significant, the observed values were in line with the direction of our hypothesis, that is, that increased automation of presence decreases the emotional engagement. The valence ($M = 0.12$) and activation ($M = 0.53$) scores for the pre-recorded video condition, suggest that emotional engagement is present in the interaction with an automated interviewer which is consistent with the literature regarding human-computer interaction. More specifically, there is research showing that people tend to treat computers as social beings and react to them as if they were humans (Nass & Moon, 2000; Reeves & Nass, 1996), and that people even more strongly respond socially to computer characters such as avatars due to their more pronounced human-like attributes (e.g., von der Pütten, 2010). The lack of significant differences between the conditions on activation and valence of affect could be related to how the interaction was perceived. The pre-recorded video interaction might have led interviewees to believe that the interaction was a live human. Perceiving the interaction as human-operated influences the social presence positively while perceiving the interaction as computer-operated negatively affects social presence (Blasovisch et al., 2002).

The Effects of a Live Versus an Automated Interviewer on the Emotional Reactions of Interviewees

We found no significant main effect of automation on the amount of emotional facial reactions. Nevertheless, joyful expressions were more prevalent when interacting with the live interviewer or live-scripted interviewer compared to the pre-recorded video interviewer. This effect, however, did not occur when interacting with the live interviewer compared to the live-scripted interviewer. The behavioral measure is perhaps a more objective measure of the interviewee's emotional reactions since no self-rated assessment of emotional reactions was included. The results provide some support to our hypothesis that increased automation of presence decreases emotional reactions. The results suggest that there are differences in a live interaction compared to a video-based interaction with regard to the amount of joyful expressions elicited. Moreover, this outcome suggests that there are no differences in a semi-scripted live interaction compared to a scripted live interaction. A possible explanation for the difference in joyful expressions but not on other emotions is that AFFDEX detected happy expressions more often. AFFDEX identifies dynamic happy expressions particularly well (91%

of the time), compared to the accuracy for other dynamic expressions: anger (49%), contempt (68%), disgust (79%), fear (1%), sadness (35%) and surprise (61%) (Stöckl et al., 2018). The reason for this could be that happiness is the most distinctly expressed emotion, whereas fear and surprise are more often confused (since they are characterized by the same markers). In line with the results for self-reported emotional engagement, the outcome for emotional reactions in the pre-recorded video condition (see Table 1) also support the existing research that people react socially to computers and human representations (Nass & Moon, 2000).

The Effects of a Live Versus an Automated Interviewer on the Interviewees' Perceived Empathy of the Interviewer

Automation of presence (i.e., less rich social cues) was found to have the greatest influence on the interviewees' perceptions of the amount of empathic responses from the interviewer. In line with our expectations, interviewees experienced least empathy in the pre-recorded video condition and most empathy in the live condition where the interviewer paraphrased or validated the interviewee. The interviewer in the live-scripted condition was also perceived as more empathetic than the interviewer in the pre-recorded-video condition. The outcomes suggest that the conditions differ in richness of social cues. This resonates with previous research that has shown that, for at least shorter interactions, media with richer social cues evoke more social presence than media with less rich social cues (e.g., Bailenson et al., 2005; Bente et al., 2008). A possible explanation for lower empathy reports could be that the pre-recorded video interview was perceived as less realistic than the live interviews. Previous studies have shown that social presence is experienced when behavioral realism is higher (von der Pütten, 2010; Pan et al., 2008; Bente et al., 2008) and that incongruity between photographic and behavioral realism yields less experience of social presence (Oh, Bailenson & Welch, 2018). Another possibility for the lower empathy reports in the pre-recorded video interview is that the pre-recorded video interview was perceived as controlled by an algorithm. Perceiving the interview as controlled by a computer has been shown to negatively impact social presence compared to when the interview is perceived as controlled by a human (Fox et al., 2015). Future studies could measure perceived agency to rule out such explanations.

In contrast to our expectation that the rigidity in the live-scripted condition would lead to fewer perceptions of empathy of the interviewer compared to the live condition, we detected no significant difference between the live and live-scripted condition, and the mean score of

empathic responses was 4.19 (on a 5-point scale) for the live interview and 3.97 for the live-scripted interview. Previous studies have shown that consistency between level of realism of visual appearance and level of realism of behavior increases social presence (Garau et al., 2003; Bailenson et al., 2005). We argue that this result suggests that the live-scripted condition, where the flexibility of the verbal responses was restricted, was perhaps not perceived as less realistic than the live condition. Increased rigidity through adherence to a script might not have led to a significant decrease in richness of social cues, in order to detect significant differences between the live and live-scripted condition. This is perhaps because social cues such as adequate eye contact, smiling, head nods and vocal pitch was still present in the live-scripted condition, compared to the pre-recorded video condition containing the same set of social cues. Together, our findings provide partial support for our hypothesis that increased automation of presence (i.e., less rich social cues) decreases perceived empathy, and suggests that a live interaction differs from a pre-recorded video interaction regarding perceptions of empathy.

With regard to the ability of the interviewer in the pre-recorded video to be empathetic, we can say that participants did not particularly perceive the interviewer in the pre-recorded video as unemphatic ($M = 3.34$ on a 5-point scale). In particular when it comes to listening skills such as paying attention, or ability to explain things clearly, the interviews did not differ significantly (see Table 1). Since this cannot be concluded statistically by traditional significance testing future studies should perform equivalence testing for statistical validity.

The Relationship between Emotional Engagement, Emotional Reactions and Perceived Empathy

We investigated the relationship between self-rated emotional engagement, emotional reactions and perceived empathy. Our hypothesis was that there would be a positive correlation between the measures. In line with our hypothesis, participants who perceived the interviewer as more empathic were found to more likely have a positive experience (i.e., greater positive change in activation of affect and more expressed joy), which might indicate that perceiving the interviewer as more empathetic contributed to a more positive experience. This outcome mirrors the research on the importance of empathy for the development of positive alliance (Nienhuis et al., 2018, Watson et al., 2014). Our results suggests that empathy has an impact on how the interaction is experienced online. This outcome also seems to reflect the above results that a pre-recorded video interaction is perceived and experienced differently (i.e., as less empathetic and

eliciting fewer joyful expressions) compared to a live interaction with a human. Contrary to what was expected, we also found a significant negative correlation between valence of affect and expressed fear. This result was likely due to AFFDEX's poor ability to correctly identify fear in only 1 % of the cases (Stöckl et al., 2018).

Limitations

As the test instructor in the pre-recorded video condition could not see the interviewee through the computer screen, while operating the interview simulation, and therefore was only able to hear them, we could not be certain that the participants had their head turned to the camera during the whole interview. If so, the AFFDEX software would not be able to analyse facial expressions since it is not applicable for different head poses. The AFFDEX data however suggest that the participants had their heads turned to the camera as the program made interpretations at each second. In each condition, the AFFDEX data was missing 1-2 times for 20-60 seconds during the interview for only a few participant. In each condition there were sometimes technical issues on our end, which sometimes resulted in rescheduling of the interview. Also, as was discussed in the Results, there were loss of AFFDEX data in the live and live-scripted condition due to a software crash. Bandwidth delay occurred at times in the live interview, which likely was due to poor internet connection. The lack of desynchronized audio and graphics in the pre-recorded video interview could possibly have influenced how realistic the interview seamed and be a potential confounding variable. It should however be added that bandwidth delay is an inherent aspect of videoconference communication and is as such a reflection of actual online communication.

Lastly, there were concerns regarding validity of the discriminant analysis to our data. Intercorrelations of our discriminator variables exceeded the level .3, in general thought to be an indication of the violation of the independence assumption (Tabachnick & Fidell, 1989). In our study this could mean that the discriminator variables did not contribute in a unique way to the group differences. Despite these concerns, requirements of complete data on the discriminator variables, knowledge of group membership, sufficient sample size, and other statistical assumptions were met, and we consider discriminant analysis to be suitable to our research questions.

Conclusions and future directions

In conclusion, our results add to the field of research aiming to examine the role of

human presence in online interventions. As automated programs are emerging in clinical settings, where people interact with a computer instead of a human, it is interesting that little is known about how automated interactions are perceived and experienced. A deeper understanding of the role of human presence is needed for development of automated internet-based self-help programs and to find ways to enhance their use by making them more engaging, since to date, adherence to these programs is low. To our knowledge, this is the first study investigating if there are differences between a live online interaction, a live-scripted interaction and a video-based interaction regarding social presence. Since the concept of social presence is defined differently in the literature, and the measures of social presence differ, our results might be difficult to compare to other similar studies. Nevertheless, our study is an important contribution to the research. The results of our study suggest that the automatization of an interviewer decreased the perceived empathy of the interviewer and decreases the emotional engagement of the interviewee in terms of expressed joy, but does not affect the self-rated engagement of the interviewee. Furthermore, increased rigidity through adherence to a script, in the live-scripted condition, did not significantly impact perceived empathy, emotional reactions, or expressed joy, compared with the live condition, rather the negative effect of automation was only found in the pre-recorded video condition. In the future, these findings should be replicated to conclude the degree of validity to the findings and more research is needed to determine if an interaction could be automatized and still be as beneficial as a live interaction.

Swedish Summary – Svensk sammanfattning

Effekterna av en live- versus automatiserad intervjuare på de intervjuades emotionella engagemang och upplevelse av empati från intervjuaren

Introduktion

Världen över lever uppskattnings 322 miljoner med en depressiv störning och 264 miljoner med en ångeststörning (World Health Organization, 2017). För att öka möjligheten för personer med psykologiska problem att få den hjälp de behöver har bland annat internetbaserad kognitiv beteendeterapi (KBT) utvecklats (Barak & Grohol, 2011). KBT-metoder över internet har visats vara lika effektiva som sedvanlig behandling öga mot öga för olika typer av psykologisk problematik (t.ex., fobier eller social ångest), både på kort och lång sikt (Andersson, 2018). Behandlingen, som går ut på att man går igenom ett program med någon form av terapeutstöd (t.ex., via e-post eller telefon), baserar sig på KBT, där man försöker öka det psykiska välmående och hjälpa individen hantera psykologiska problem genom att förändra tankar och beteenden (INSERM, 2004; Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012).

Det mänskliga stödet är centralt för ett gott behandlingsresultat i internet-KBT (Andersson, 2009), men det finns även studier som påvisar att interventioner utan mänsklig närvaro kan vara effektiva, till och med på lång sikt (e.g., Griffiths, Farrer & Christensen, 2010, Vedaa, 2019). Fördelarna med självhjälpsprogram jämfört med vägledd internet-KBT är att de lättare kan distribueras i en större skala, är billigare och också kunde erbjudas till personer i väntan på behandling (Leykin, Muñoz, Contreras, & Latham, 2014). Fastän internetbaserade interventioner är ett lovande alternativ för behandling av mindre komplexa psykologiska problem, så borde internetbaserade interventioner inte ses som ett substitut till traditionell terapi. Det här beror på att internetbaserade behandlingar inte lämpar sig för alla eller för svårare psykologisk problematik (Andersson & Titov, 2014).

Den största nackdelen med självhjälpsprogram är att de inte är tillräckligt engagerande vilket leder till att många inte slutför en påbörjad behandling (Christensen, Griffiths, & Farrer, 2009). Man har provat på att inkludera simulerat stöd i behandlingen, i form av interaktion med en avatar (dvs. en datoriserad representation av en psykolog), vilket har visat sig påverka positivt på relationen med programmet (Pinto et al., 2013, 2016; Rehm et al., 2016). I en studie av Pinto et al., (2013, 2016) fann man att deltagare som interagerade med avataren upplevde en kontakt

med avataren och rapporterade färre depressionssymptom efter behandlingen jämfört med de deltagare som inte fick simulerat stöd (Pinto et al., 2013, 2016; Rehm et al., 2016). Pinto et al. (2016) spekulerade att resultatet kan bero på interaktionen med avataren. I traditionell behandling predicerar den terapeutiska relationen behandlingsresultatet (Norcross & Lambert, 2018) och det är trolig att också i den internetbaserade interventioner är den terapeutisk relation viktig. Ovannämnda forskningsfynd kan ändå betyda att närvaro av en verklig människa inte alltid är nödvändigt för att en intervention ska vara verksam eller för att simulerad närvaro kan vara lika effektiv som mänskligt stöd. För att få en djupare förståelse för den mänskliga närvarons roll i en internetbaserad behandling och vilken mängd och typ av närvaro som krävs för att en behandling ska vara effektiv, så undersökte vi hur en simulerad eller automatiserad mänsklig närvaro uppfattas i en onlinemiljö. I en 10–15-minutersintervju undersökte vi hur automatisering av en mänsklig närvaro påverkar faktorer som är centrala för den terapeutiska relationen, dvs. deltagarnas emotionella engagemang och deras upplevelse av empati från intervjuaren. Vi skapade tre betingelser vilka innefattade en vanlig intervju (live), en mera rigid intervju (skriptad live) och en helt automatiserad intervju (inspelad video). I alla tre betingelser mätte vi deltagarnas reaktioner och deras upplevelse av empati från intervjuaren.

Den terapeutiska alliansen, det emotionella engagemanget och upplevd empati i kliniska interaktioner

I traditionell behandling är en stark relation mellan klienten och terapeuten central för behandlingsresultatet (Norcross & Lambert, 2018). En grundläggande del av relationen och behandlingsresultatet är terapeutens och klientens emotionella engagemang och terapeutens empatiska förmåga i relationen har även visats predicera förändring hos klienten (Elliot et al., 2018; Norcross & Lambert, 2018). Förutom detta bidrar empati också till att bygga en positiv allians mellan terapeuten och klienten (Nienhuis et al., 2018, Watson et al., 2014) och påverkar samspelet i relationen och klientens engagemang positivt (Watson et al., 2014). Fastän alliansen i vägledda internetinterventioner påvisats vara lika stark som i en behandling öga mot öga (Knaevelsrud and Maercker, 2006; Preschl et al., 2011; Wagner et al., 2012) är terapeutens roll i internetbaserade behandlingar mer otydlig eftersom relationen inte lika ofta predicerar ett positivt behandlingsresultat (Berger, 2017; Andersson, 2018). Forskning har visat att en klient också kan skapa en allians med ett datorprogram (Bickmore, Gruber, & Picard, 2005; Bickmore, Caruso, Clough-Gorr, & Heeren, 2005; Ormrod et al, 2010). En allians med ett datorprogram är

möjlig eftersom att mänskor tenderar att behandla datorer som sociala varelser. Genom olika studier har Reeves, Nass med sitt forskningsteam visat att människor tenderar regera på sociala signaler oavsett om miljön är verklig eller datorgenererad. Detta paradigm kallas CASA (Computers as Social Actors) (Reeves & Nass, 1996, Nass & Moon, 2000). Fastän man vet att mänskor interagerar med datorer så som man skulle interagera med mänskor så vet man litet om interaktionen med en dator upplevs som empatisk och emotionellt engagerande. I takt med att människa-dator-kommunikation blir vanligare i kliniska kontexter vore det viktigt att undersöka dessa variabler.

Social närvaro i kliniska interaktioner

Att den andra personen (verkliga eller mänskorepresentationen) upplevs som närvarande, kallas social närvaro och är en viktig aspekt av människa-dator-interaktionen (Lee, 2004). Mediet som används i kommunikationen verkar vara viktigt för upplevelsen av social närvaro. I en studie av Bente et al. (2008) fann man att textbaserad kommunikation väckte mindre social närvaro jämfört med medier med mera nyanserade sociala signaler (e.g., video, audio eller avatar). Man har visat att tiden också påverkar upplevelsen av social närvaro så att också medier med mindre nyanserade sociala signaler kan väcka lika mycket social närvaro eftersom mänskor anpassar sig till kommunikationsformen och utvecklar nya kommunikationsstrategier (e.g., Ramirez et al., 2002; Antheunis et al., 2010, Walther, 1992).

Också hur social närvaro representeras visuellt och hur realistiskt representationen beter sig påverkar hur den sociala närvaron upplevs. Ju mera den visuella representationen beter sig som en människa (dvs. beteenderealism), desto positivare inverkan har den på upplevelsen av social närvaro (e.g., Bente et al., 2008; Pan et al., 2008; Pütten et al., 2010). Exempelvis i en studie av Pütten et al. (2010) fann man att en avatar som nickade väckte mera social närvaro hos deltagarna jämfört med en avatar som inte nickade. Man har också förevisat liknande effekter för beteenden så som ömsesidig ögonkontakt (Bente et al., 2008) och en agent som rodnar efter att ha misslyckats i en presentation (Pan et al., 2008). Beteenderealism verkar påverkas av hur representationen ser ut. Det här betyder att ju mer lik en människa representationen är, desto mer trovärdigt måste beteendet också vara för en starkare upplevelse av social närvaro (Bailenson et al., 2005; Garau et al., 2003). Det finns ingen konsensus kring hur man mäter social närvaro, eftersom olika forskare använder olika definitioner av social närvaro (Bailenson et al., 2005). Social närvaro har både operationaliserats som graden till vilken en person upplever att den

andra personen är närvarande och som en social reaktion (Bailenson et al., 2005). I föreliggande studie var det huvudsakliga måttet för social närvaro graden av upplevd närvaro och empati från intervjuaren. Vi mätte även social närvaro som sociala reaktioner. Detta gjorde vi genom att mäta deltagarnas självskattade emotionella engagemang och analysera deras ansiktsreaktioner med ett mjukvaruprogram.

Studiens syfte och hypoteser

Det huvudsakliga målet med studien var att undersöka om det fanns skillnader mellan en liveinteraktion, en skriptad liveinteraktion och en inspelad videointeraktion. I enlighet med CASA-paradigmet (Nass & Moon, 2000), och eftersom betingelserna i hög grad vara lika, så förväntade vi oss att deltagarna i den inspelade videointervjun skulle visa liknande emotionellt engagemang som i liveintervjun. Vi antog också att intervjuaren i liveintervjun skulle förmedla mera nyanserade sociala signaler än den automatiserade intervjuaren (dvs. intervjuaren i den inspelade videointervjun) och därför också skulle väcka mera social närvaro. Vi tänkte oss att rigiditeten i den inspelade videointervjun kan påverka upplevelsen av närvaro negativt eftersom interaktionen inte är lika spontan och naturlig som i liveinteraktionen. Vi förväntade oss alltså att betingelserna skulle väcka liknande reaktioner, men ifall rigiditeten påverkar upplevelsen av social närvaro så borde vi kunna hitta skillnader mellan betingelserna. Vi undersökte även om upplevd empati, självskattat emotionellt engagemang och emotionella reaktioner korrelerar positivt eftersom man visat att en terapeuts empatiska förmåga bidrar till att skapa en god allians och därmed även påverkar det emotionella engagemanget (Nienhuis et al., 2018, Watson et al., 2014). På basis av tidigare nämnd forskning testade vi därför följande hypoteser:

Hypotes 1: Ökad automatiseringsgrad av närvaro (dvs. mindre nyanserade sociala signaler) kommer att minska upplevt emotionellt engagemang, emotionella reaktioner och upplevd empati så att livebetingelsen väcker mest upplevt emotionellt engagemang, emotionella reaktioner och upplevd empati, den skriptade livebetingelsen näst mest och den inspelade videobetingelsen minst. Den alternativa hypotesen vara att det inte föreligger några skillnader mellan betingelserna.

Hypotes 2: Upplevd empati korrelerar positivt med upplevt emotionellt engagemang och emotionella reaktioner.

Metod

Studien erhöll etiskt tillstånd våren 2017 från den forskningsetiska nämnden vid fakulteten för humaniora, psykologi och teologi vid Åbo Akademi.

Deltagarna rekryterades via universitetet och sociala medier. Vårt sampel bestod av 75 universitetsstuderande, av vilka 49 var kvinnor, 25 män och en deltagare som inte uppgav information om könstillhörighet.

För att studera ovannämnda frågeställning genomförde vi ett experiment med mellanindivid-design där automatiseringsgrad av närvaro (dvs. mindre nyanserade sociala signaler) operationaliserades utifrån typ av närvaro och flexibiliteten i intervjuarens svarsalternativ (i.e., live versus skriptad live vs. inspelad video). Studien innefattade följande tre betingelser: (1) en intervju live där deltagarna såg en psykologistuderande på en datorskärm och blev intervjuade muntligt enligt ett skript. Intervjuaren i livebetingelsen (i.e., psykologistuderanden) kunde ge ett verbalt och icke-verbalt parafraserande eller validerande svar mellan frågorna. (2) En intervju live där deltagaren såg en psykologistuderande på en datorskärm och intervjuades på samma vis som i betingelse 1. Här kunde intervjuaren däremot enbart ge ett skriptat svar mellan frågorna (se Figur 1). (3) En videobaserad intervju där intervjuaren ersatts med inspelade videoklipp av intervjuaren. Svaren var skriptade som i betingelse 2. Replikerna och frågorna följde samma ordning som i betingelse 1 och 2. Testledaren (dvs. intervjuaren) i experimentet styrde videoklippen.

Vi använde det validerade självskattningsformuläret Swedish Core and Affect Scale (SCAS) (Västfjäll et al., 2002; Västfjäll & Gärling, 2007) för att mäta upplevt emotionellt engagemang före och efter intervjun. Mjukvaruprogrammet AFFDEX Software Development Kit (SDK) 2.0. användes för att mäta deltagarnas emotionella reaktioner under intervjun. Genom att kartlägga ansiktsrörelser kan programmet på basis av ett kodningssystem identifiera emotionerna ilska, sorgsenhet, avsky, glädje, förvåning och förakt (Ekman & Friesen, 1976). Vi mätte upplevd empati med den svenska versionen av det validerade frågeformuläret Consultation and Relational Empathy (CARE) measure (Mercer et al., 2004, Swedish translation; Ahlforn, Horwitz, & Osika, 2017). Vi samlade även in demografisk information om deltagarnas kön, ålder och erfarenhet av psykoterapi eller annan psykologisk behandling.

I alla betingelserna genomförde vi intervjuerna på en iMac (24-inch, 1920x1200, 2009) och en HP Compac 8200 Elite MT PC. Webbsidan Doxy.me - © Doxy.me, LLC 2014 användes för liveintervjuerna. Till den inspelade intervjun filmades totalt 27 videoklipp (14 intervjufrågor,

4 inledande repliker, en åsnebrygga, 6 skriptade svar, ett avslutande svar, och ett bakgrundsklipp) med en digitalkamera (Canon EOS 1300D; lens: Canon EF 50mm f/1.8). För att skapa en smidig simulering av den skriptade livebetingelsen editerades sedan videoklippen med Resolume Arena 5. Intervjuaren satt i ett annat rum och styrde simuleringen.

Före huvudexperimentet gjorde vi ett pilotförsök med ett sampel på tio deltagare för att testa intervjufrågornas gångbarhet. Vi utformade egna intervjufrågor utgående från teman som vanligen behandlas i kliniska intervjuer (Sommers-Flanagan & Shaw, 2017). Vi gjorde inga ändringar efter pilotförsöket. Intervjuerna utfördes på psykologiinstitutionen vid Åbo Akademi. För att kunna testa vår hypotes delgavs deltagarna vid rekryteringen varken information om studiens syfte eller hypoteser, men fick veta att temat för den 10-15-minuterlånga intervjun handlade om deras vardag och hälsa (se Appendix). Studien var dubbelblind eftersom varken deltagarna eller assistenten som gav deltagarna instruktioner kände till forskningsfrågan, hypoteserna eller studiens upplägg. Deltagarna blev instruerade att undvika att byta sittställning eller svänga på huvudet under intervjun och för att inte höra intervjuaren tala i rummet bredvid använde deltagarna hörlurar. Efter intervjun fick deltagarna debriefing gällande studiens syfte och fick information om stadens psykologtjänster.

Resultat

Vi fann inga signifikanta skillnader mellan grupperna gällande gruppstorlek, könsfördelning, depressionssymptom och tidigare erfarenhet av psykologisk behandling, och därför beaktades dessa variabler inte i huvudanalyserna. Innan huvudanalyserna kollade vi för hur många värden som saknas i data och fann att 49 (9 %) av observationerna i AFFDEX-data saknades. Orsaken till de saknade värdena var en krasch i mjukvaruprogrammet i fall då programmet startats innan inloggning till webbsidan Doxy.me - © Doxy.me, LLC 2014. Eftersom Doxy.me bara användes i liveintervjun och den skriptade liveintervjun var datakraschen enbart möjlig i dessa betingelser, vilket indikerar att data var kopplad till gruppvariablen och därmed inte saknades slumpmässigt (Little & Rubin, 2015). Vi hanterade de saknade värdena med hjälp av multipel imputering med variabelns medelvärde som imputationsvärde. Endast 1 % av värdena från bakgrundsformuläret och 2 % av värdena från SCAS-formuläret saknades. I analyserna exkluderade vi de observationer som innehöll saknade värden.

För att undersöka om det fanns någon skillnad mellan betingelserna gällande deltagarnas

upplevda emotionella engagemang före och efter intervjun genomförde vi en variansanalys för upprepade mätningar. Vi hittade ingen signifikant skillnad mellan betingelserna för varken känsloreaktionens valens, $F(2,75) = 0.63$, $p = .939$, eller aktiveringsgrad, $F(2,75) = 2.88$, $p = .063$. Resultaten tyder på att liknande emotionellt engagemang upplevts i de olika betingelserna, men att det inte fanns någon skillnad mellan live-, den skriptade live- och den inspelade betingelsen.

Med multivariat variansanalys, undersökte vi om det fanns någon skillnad i emotionella reaktioner mellan betingelserna. Vi hittade ingen signifikant huvudeffekt av ökad automatiseringsgrad på mängd emotionella reaktioner, $F(14,134) = 1.153$, $p = .319$, $V = .215$. Däremot fann vi att ökad automatiseringsgrad minskade uttryckt glädje signifikant, $F(2,75) = 5.49$, $p < .01$ partial $\eta^2 = .132$. I liveintervjun och den skriptade liveintervjun uttryckte deltagarna mindre glädje än i den inspelade videointervjun. Det fanns ingen skillnad i emotionella reaktioner mellan liveintervjun och den skriptade liveintervjun.

Vi undersökte om deltagarna i live-, den skriptade, och den inspelade betingelsen upplevde intervjuaren olika och fann en signifikant huvudeffekt av ökad automatiseringsgrad på upplevd empati, $V = .510$; $F(20,128) = 2.191$, $p = .005$; (partial) $\eta^2 = .255$. Med en diskriminantanalys analyserade vi sedan hur gruppskillnaderna såg ut och vilka variabler på CARE-måttet som bäst kunde förklara skillnaden. Med en diskriminantanalys kan man predicera hur många funktioner som behövs för att beskriva sambandet mellan de oberoende variablerna (dvs. CARE-variablerna) och en kategorisk variabel (dvs. grupptillhörighet). Vi fann att skillnaden bäst förklarades av en diskriminantfunktion, $L = 0.594$, $\chi^2(20) = 35.892$, $p = .001$, som utgjorde 89,2 % av variansen, kanonisk $R^2 = .37$. Hur bra intervjuaren var på att få den intervjuade att känna sig avslappnad ($r = .74$) och ha en positiv hållning ($r = .71$) korrelerade mest med diskriminantfunktionen jämfört med de andra CARE-variablerna. Analysresultaten tyder på att deltagarna i den inspelade videointervjun (r = -1.048) upplevde intervjuaren som mindre positiv och sämre på att få en att känna sig avslappnad jämfört med livebetingelsen (r = .539) och den skriptade livebetingelsen (r = .509). Det förelåg ingen skillnad mellan den skriptade liveintervjun och liveintervjun. I jämförelse med slumpmässig klassificering (33,33%) kunde diskriminantfunktionen utgående från värden på CARE-variablerna predicera grupptillhörighet med en noggrannhet på 61,33% och med korsvalidering med en noggrannhet på 52,00%.

Till sist undersökte vi om det fanns en positiv korrelation mellan upplevd empati, självskattat emotionellt engagemang och emotionella reaktioner. Korrelationsanalysen visade att upplevd empati korrelerade positivt med känsloreaktionens aktiveringsgrad, $r = .32$; $p = .006$; $R^2 = .10$, och glada ansiktsuttryck, $r = .29$; $p = .01$; $R^2 = .08$. På basis av resultaten kan korrelationens riktning inte bestämmas, men resultaten tyder på att de deltagare som upplevde intervjuaren som mer empatisk också upplevde intervjun som mer positiv (dvs. uttryckte mer glädje och upplevde sig aktiveras mera emotionellt). Vi fann även en negativ korrelation mellan känsloreaktionens valens och uttryckt rädsla, $r = -.25$; $p = .037$. Den negativa korrelationen kan bero på felaktig identifikation eftersom AFFDEX identifierar rädsla med en noggrannhet på föga 1 % (Stöckl et al., 2018). Vi fann även inbördes korrelationer mellan olika emotioner vilket kan tyda på att deltagare som uttryckte en viss emotion mer sannolikt också uttryckte andra emotioner, dvs. var mera emotionellt engagerad i intervjun.

Diskussion

Syftet med studien var att undersöka om det fanns skillnader mellan en liveinteraktion, en skriptad liveinteraktion och en inspelad videointeraktion. I enlighet med CASA-paradigmet (Nass & Moon, 2000), och eftersom betingelserna i hög grad vara lika, så förväntade vi oss att deltagarna i den inspelade videointervjun skulle visa liknande reaktioner och emotionellt engagemang som i liveintervjun. Vi antog dessutom att intervjuaren i liveintervjun skulle förmedla mera nyanserade sociala signaler och därmed också väcka mera social närvaro än den automatiserade intervjuaren (dvs. intervjuaren i den inspelade videointervjun). Vi tänkte oss att rigiditeten i den inspelade videointervjun kan påverka upplevelsen av närvaro negativt eftersom interaktionen inte är lika spontan och naturlig som i liveinteraktionen. Vi förväntade oss alltså att betingelserna skulle väcka liknande reaktioner, men ifall rigiditeten påverkar upplevelsen av social närvaro så förväntade vi oss kunna hitta skillnader mellan betingelserna. Vi undersökte även om upplevd empati, självskattat emotionellt engagemang och emotionella reaktioner korrelerar positivt eftersom man visat att en terapeuts empatiska förmåga bidrar till att skapa en god allians och därmed även sannolikt påverkar det emotionella engagemanget (Nienhuis et al., 2018, Watson et al., 2014).

Vi fann inte stöd för hypotesen att ökad automatiseringsgrad av närvaro (dvs. mindre nyanserade sociala signaler) minskar upplevt emotionellt engagemang (se Table 1). Känsloreaktionernas valens och aktiveringsgrad i den inspelade videointervjun (se Table 1) tyder

på att deltagarna engagerade sig emotionellt i intervjun, vilket överensstämmer med tidigare forskningsresultat som visat att mänskor tenderar interagera med datorer och avatarer på samma sätt som med mänskor (Nass & Moon, 2000; Reeves & Nass, 1996; von der Pütten, 2010). Huruvida deltagarna i den inspelade videobetingelsen uppfattat intervjuaren som en verklig person eller styrd av en dator kan ha påverkat resultaten. Tidigare forskning har visat att om interaktionen uppfattas som människostyrd upplevs den andra som mera socialt närvarande (Blasovisch et al., 2002).

Vi fann heller inte stöd för att ökad automatiseringsgrad (dvs. mindre nyanserade sociala signaler) påverkar mängd emotionella reaktioner negativt. Däremot visade deltagare oftare glada ansiktsuttryck när de interagerade med intervjuaren i livebetingelsen eller den skriptade livebetingelsen jämfört med då intervjuaren ersatts av inspelade videoklipp. Vi fann ingen skillnad mellan livebetingelsen och den skriptade livebetingelsen. Det fanns ingen skillnad i mängd glada ansiktsuttryck mellan livebetingelsen och den skriptade livebetingelsen. Resultaten ger något stöd för vår hypotes att ökad automatiseringsgrad av närvaro (dvs. mindre nyanserade sociala signaler) minskar emotionella reaktioner och indikerar att det finns skillnader mellan en liveinteraktion och en videobaserad interaktion med avseende på uttryckt glädje. Därtill tyder resultaten på att det inte finns någon skillnad mellan en semiskriptad liveintervju och en skriptad liveintervju. En möjlig orsak till att det enbart fanns skillnader i glada ansiktsuttryck kan bero på att AFFDEX med större noggrannhet identifierar ett glatt ansiktsuttryck (91% av tiden) jämfört med ilska (49%), förakt (68%), avsky (79%), rädsla (1%), sorgsenhet (35%) och förvåning (61%) (Stöckl et al., 2018). Att olika emotioner visades i den inspelade videointeraktionen (se Table 1) stöder tidigare forskningsresultat om att mänskor reagerar socialt på datorer och datorrepresentationer (Nass & Moon, 2000; von der Pütten, 2010).

Däremot hittade vi en signifikant effekt av automatiseringsgrad (dvs. mindre nyanserade sociala signaler) på deltagarnas upplevelse av empati från intervjuaren. Enligt våra förväntningar så rapporterade deltagarna i den inspelade liveintervjun minst upplevd empati och mest upplevd empati i liveintervjun där intervjuaren gett parafraserande eller validerande responser. Intervjuaren i den skriptade liveibetingelsen upplevdes också som mer empatisk jämfört med intervjuaren i den inspelade videobetingelsen. Våra resultat tyder på att betingelserna skiljde sig i hur nyanserat det sociala innehållet var vilket också berör tidigare forskningsfynd som visat att medier med mera nyanserade sociala signaler stärker upplevelsen av social närvaro, åtminstone i

kortare interaktioner (e.g., Bailenson et al., 2005; Bente et al., 2008). Eventuellt kan vårt resultat bero på att den inspelade videointervjun upplevts som mindre realistisk än liveintervjuerna. Tidigare forskning har visat att mer realistiskt beteende samt att om beteendet är lika realistiskt som utseendet ökar upplevelsen av social närvaro (von der Pütten, 2010; Pan et al., 2008; Bente et al., 2008; Oh, Bailenson & Welch, 2018). En annan möjlig orsak till utfallet kan vara att den inspelade videointervjun upplevts som styrd av en algoritm. I tidigare studier har man visat att upplevelsen av social närvaro påverkas negativt om interaktionen upplevs vara styrd av en dator jämfört med om den upplevs vara styrd av en människa (Fox et al., 2015).

I motsats till vår hypotes om att rigiditeten i den skriptade livebetingelsen skulle leda till att man upplever mindre empati från intervjuaren jämfört med intervjuaren i livebetingelsen så hittade vi inga signifikanta skillnader mellan livebetingelsen och den skriptade livebetingelsen. Ttidigare studier har visat att en starkare upplevelse av social närvaro väcks då utseendet är lika realistiskt som beteendet (Garau et al., 2003; Bailenson et al., 2005). Våra resultat kan tyda på att livebetingelsen, med skriptade svar, inte uppfattats som mindre realistisk än liveintervjun med validerande och parafraserande svar. Eventuellt ledde ökad rigiditet (dvs. att intervjuaren höll sig till skriptade svar) inte till signifikant mindre nyanserade sociala signaler för att kunna upptäcka skillnader mellan livebetingelsen och den skriptade livebetingelsen. Det här kan bero på att intervjuaren i den skriptade liveintervjun fortfarande kunde interagera mer naturligt (t.ex., hålla adekvat ögonkontakt, le när det var lämpligt att le), jämfört med den inspelade videointervjun vars sociala signaler alltid var likadana.

Vi undersökte även sambandet mellan självskattat emotionellt engagemang, emotionella reaktioner och upplev empati från intervjuaren. Vår hypotes var att måtten skulle korrelera positivt. I enlighet med vår hypotes fann vi att deltagare som tillskrivit intervjuaren mer empatisk förmåga oftare upplevde intervjun mer positivt (dvs., upplevde sig aktiveras mer emotionellt och uttryckte mer glada ansiktsuttryck), vilket kan indikera att upplevelse av empati bidrog till en mer positiv erfarenhet. Resultaten verkar även reflektera våra andra resultat att den inspelade videointeraktionen uppfattas och upplevs annorlunda än en liveinteraktion med en människa. I motsats till vad som förväntades hittade vi en signifikant negativ korrelation mellan känsloreaktionens valens och arga ansiktsuttryck. Det här resultatet berodde sannolikt på att AFFDEX enbart i 1 % av fallen gjort en korrekt identifiering av rädsla (Stöckl et al., 2018).

Begränsningar med den föreliggande studien var att experimentledaren som styrde den inspelade videobetingelsen inte kunde se deltagaren via datorskärmen. Detta medförde att vi inte kunde vara säkra på att deltagarna hade huvudet vänt mot kameran under intervjun. AFFDEX-data antyder ändå att högst fem deltagare vänt huvudet från kameran högst två gånger i högst en minur under intervjun. Dålig nätverksanslutning ledde ibland till att ljud- eller videokvaliteten var hackig eller fördröjd. Att videon fördröjs eller låser sig är vanliga problem i internetkommunikation och kan sålunda ha påverkat ifall den inspelade videointervjun upplevts som realistis. I diskriminantanalysen bröts antagandet om oberoende observationer vilket kan innebära att diskriminantfunktionerna inte på ett unikt sätt förklarade gruppskillnaderna (Tabachnick & Fidell, 1989).

Såvitt vi vet är vår studie den första som undersökt ifall det föreligger skillnader mellan en semiskriptad liveinteraktion, en skriptad liveinteraktion och en inspelad videointeraktion gällande upplevelse av social närvaro. Våra resultat kan vara svåra att jämföra med liknande studier eftersom begreppet social närvaro definierats på olika sätt och det också finns varierande mått av social närvaro. Trots det är vår studie ett viktigt bidrag till forskningen som undersöker betydelsen av mänsklig närvaro i internetbaserade interventioner. Våra resultat tyder på att en automatiserad intervjuare minskade upplevelsen av empati från intervjuaren och deltagarnas emotionella engagemang beträffande uttryckt glädje, men påverkade inte deltagarnas självskattade emotionella engagemang. Dessutom påverkade ökad rigiditet (dvs. att hålla sig till skriptade svar) inte upplevd empati, emotionella reaktioner eller uttryckt glädje jämfört med livebetingelsen. Med andra ord fanns det endast en effekt av automatisering i den inspelade videobetingelsen. Resultaten borde replikeras för att bedöma resultatens giltighet och mer forskning behövs för att kunna bedöma huruvida en interaktion kan automatiseras och vara lika verksam som en liveinteraktion.

References

Ahlforn, K.C., Horwitz, E.B., & Osika, W. (2017). A Swedish version of the Consultation and Relational Empathy (CARE) measure. *Scandinavian Journal of Primary Health Care, 35*(3), 286-292. doi: 10.1080/02813432.2017.1358853

Andersson, G. (2018). Internet interventions: past, present and future. *Internet Interventions, 12,* 181-188. doi: 10.1016/j.invent.2018.03.008

Andersson, G. (2009). Using the Internet to provide cognitive behavioural therapy. *Behaviour Research and Therapy, 47,* 175-180. doi: 10.1016/j.brat.2009.01.010

Andersson, G., & Titov, N. (2014). Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry, 13*(1), 4-11. doi: 10.1002/wps.20083

Antheunis, M. L., Valkenburg, P. M., & Peter, J. (2010). Getting acquainted through social network sites: testing a model of online uncertainty reduction and social attraction. *Computers in human behavior, 26*, 100–109. doi: 10.1016/j.chb.2009.07.005

Bailenson, J.N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., & Blasovich, J. (2005). The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence, 14*(1)*,* 379-393. doi: 10.1162/105474605774785235

Barak, A., & Grohol, J.M. (2011). Current and future trends in internet-supported mental helath interventions. *Journal of Technology in Human Services, 29*(3), 155-196. doi: 10.1080/15228835.2011.616939

Bente, G., Rüggenberg, S., Krämer, N.C., & Eschenburg, F. (2008). Avatar-mediated networking: Increasing social presence and interpersonal trust in net-based collaboration. *Human Communication Research, 34*(2), 287-318. doi: 10.1111/j.1468-2958.2008.00322.x

Berger, T. (2017). The therapeutic alliance in internet interventions: a narrative review and suggestions for future research, *Psychotherapy Research, 27*(5), 511-524. doi: 10.1080/10503307

Bickmore, T., Caruso, L., Clough-Gorr, K., & Hereen, T. (2005). "It's just like you talk to a friend" Relational agents for older adults. *Interacting with Computers, 17*(6), 711-735. doi: 10.1016/j.intcom.2005.09.002

Bickmore, T., Gruber, A., & Picard, R. (2005). Establishing the computer-patient working alliance in automated health behavior change interventions. *Patient Education and Counseling, 59*(1), 21-30. doi: 10.1016/j.pec.2004.09.008

Blasovich, J., Loomis, J., Beall, A.C., Swinth, K.R., Hoyt, C.L., & Bailenson, J.N. (2002). Immersive virtual environment technology as a methological tool for social psychology. *Psychological Inquiry, 13*(2), 103-124. doi: 10.1207/S15327965PLI1302_01

Brown, M., & Wicker, L. (2000). Discriminant analysis. In H. Tinsley, & S. Brown (Eds.), *Applied multivariate statistics and mathematical modeling* (pp. 209-235). doi: 10.1016/B978-0-12-691360-6.X5000-9

Bylsma, L.M., Morris, B.H, & Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clinical Psychology Review, 28*(4), 676-691. doi: 10.1016/j.cpr.2007.10.001

Christensen, H., Griffiths, K.M., & Farrer, L. (2009). Adherence in internet interventions for depression and anxiety. *Journal of Medical Internet Research, 11*(2), e13. doi: 10.2196/jmir.1194

Cuijpers, P., Donker, T., Johansson, R., Mohr, D.C., van Straten, A., & Andersson, G. (2011) Self-guided psychological treatment for depressive symptoms: a meta-analysis, *PLoS ONE 6*(6): e21274. doi:10.1371/journal.pone.0021274

Ekman, P., & Friesen, W.V. (1976). Measuring facial movement. *Environmental Psychology and nonverbal Behavior, 1*(1), 56-75. doi: 10.1007/BF01115465

Elliott, R., Bohart, A.C., Watson, J.C., & Murphy, D. (2018). Therapist Empathy and Client Outcome: An Updated Meta-analysis. *Psychotherapy, 55*(4), 399-410. doi: 10.1037/pst0000175

Fox, J., Ahn , S.J., Janssen, J.H., Yeykelis, L., Segovia, K.Y., & Bailenson, J.N. (2015). Avatars versus agents: A meta-analysis quantifying the effect of agency on social influence, human-computer interaction, *30*(5), 401-432. doi: 10.1080/07370024.2014.921494

Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., & Sasse, M.A. (2003). The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, *USA*, *5*(1), 529-536. doi: 10.1145/642611.642703

Griffiths, K.M., Farrer, L., & Christensen, H. (2010). The efficacy of internet interventions for depression and anxiety disorders: a review of randomised controlled trials [Supplement material]. *The Medical Journal of Australia, 192*(11), S4-S11. Retrieved from https://www.mja.com.au/journal/2010/192/11/efficacy-internet-interventions-depression-and-anxiety-disorders-review

Harvey, A.G., & Gumport, N.B. (2015). Evidence-based psychological treatments for mental disorders: Modifiable barriers to access and possible solutions. *Behaviour Research and Therapy, 68,* 1-12. doi: 10.1016/j.brat.2015.02.004

Hofmann, S.G., Asnaani, A., Imke, I.J.J., Sawyer, A.T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research, 36*(5), 427-440. doi: 10.1007/s10608-012-9476-1

Holst, A., Nejati, S., Björkelund, C., Eriksson, M.C., Hange, D., Kivi, M.,..., Petersson, E.L. (2017). Patients' experiences of a computerised sef-help program for treating depression – a qualitative study of Internet mediated cognitive behavioural therapy in primary care. *Scandinavian Journal of primary Health Care, 35*(1), 46-53. doi: 10.1080/02813432.2017.1288813

INSERM Collective Expertise Centre, French National Institute of Health and Medical Research. (2004). *Psychotherapy: three approaches evaluated.* Retrieved from://www.ncbi.nlm.nih.gov/books/NBK7127/

iMotions (2016). *Facial Expression Analysis: The definitive guide*. Retrieved from https://imotions.com/facialexpression-guide-ebook/

Kersten, P., White, P., & Tennant, A. (2011). The consultation and relational empathy measure: An investigation of its scaling structure. *Disability and Rehabilitation, 34*(6), 503-509. doi: 10.3109/09638288.2011.610493

Knaevelsrud, C., & Maercker, M.D. (2006). Does the quality of the working alliance predict treatment outcome in online psychotherpy for traumatized patiens?. *Journal of Medical Internet Research, 8*(4): e31. doi: 10.2196/jmir.8.4.e31

Kroenke, K., Spitzer, R.L., & Williams, J.B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606-613. doi: 10.1046/j.1525-1497.2001.016009606.x

Lee, K.M. (2004). Presence, explicated. *Communication Theory, 14*(1), 27-50. doi: 10.1111/j.1468-2885.2004.tb00302.x

Leykin., Y., Muñoz R.F., Contreras, O., Latham M.D. (2014). Results from a trial of an unsupported internet intervention for depressive symptoms. *Internet Interventions, 1,* 175–181. doi: 10.1016/j.invent.2014.09.002

Lillevoll, K.R., Vangberg, H.C.B., Griffiths, K.M., Waterloo, K., Eisemann, M.R. (2014). Uptake and adherence of self-directed internet-based mental health intervention with tailored e-mail

reminders in senior high schools in Norway. *BMC Psychiatry, 14*(1): 14. doi: 10.1186/1471-244X-14-14

Little, R.J., & Rubin, D.B. (2015). Missing data. In J.D. Wright (2nd ed.), *International Encyclopedia of the Social and Behavioral Sciences, Vol. 2,* 602-607. *doi:* 10.1016/B978-0-08-097086-8.42082-9.

Mauss, I.B., Levenson, R.W., McCarter, L., Wilhelm, F.H., & Gross, J.J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion, 5*(2), 175-190. doi: 10.1037/1528-3542.5.2.175

Meiselman, H. L. (2016). Emotion measurement. Cambridge, UK: Woodhead.

Mercer, S.W., Maxwell, M., Heaney, D., & Watt, G.C. (2004). The consultation and relational empathy (CARE) measure: development and preliminary validation and reliability of an empathy-based consultation process measure. *Family Practice, 21*(6), 699-705. doi: 10.1093/fampra/cmh621

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81-103. doi: 10.1111/0022-4537.00153

Nienhuis, J.B., Owen, J., Valentine, J.C., Winkeljohn Black, S., Halford, T.C., Parazak, S.E., Budge, S., Hilsenroth, M. (2016). Therapeutic alliance, empathy, and genuineness in individual adult psychothereapy: a meta-analytic review. *Psychotherapy Research, 28*(4), 593-605. doi: 10.1080/10503307.2016.1204023

Norcross, J.C., & Lambert, M.J. (2018). Psychotherapy Relationships That Work III. *Psychotherapy, 55*(4), 303-315. doi: 10.1037/pst0000193

Oh, C.S., Bailenson, J.N., & Welch, G.F. (2018). A systematic review of social presence: definition antecedents, and implications. *Frontiers in Robotics and Al, 5*(14), 1-34. doi: 10.3389/frobt.2018.00114

Omrod, J., Kennedy, L., Scott, J., Cavanagh, K. (2010). Computerized cognitive behavioral therapy in an adult mental helath service: a pilot study of outcomes and alliance, *Cognitive Behaviour Therapy, 39*(3), 188-192. doi: 10.1080/16506071003675614

Pan, X., Gillies, M., & Slater, M. (October, 2008). *The impact of avatar blushing on the duration of interaction between a real and virtual person*. Paper presented at the meeting of International Society for Presence Research and Human Technology Lab, Padova, Italy.

Pinto, M.D, Hickman, R.L, Clochesy, J., Buchner, M. (2013). Avatar-based depression self-management technology: promising approach to improve depressive symptoms among young adults. *Applied Nursing Research, 26*(1), 45-48. doi: 10.1016/j.apnr.2012.08.003

Pinto, M.D., Greenblatt, A.M., Hickman, R.L., Rice, H.M., Thomas, T.L., Clochesy, J.M. (2016). Assessing the critical parameters of eSMART-MH: a promising avatar-based digital therapeutic intervention to reduce depressive symptoms. *Perspectives in Psychiatric Care, 52*(3), 157–167. doi: 10.1111/ppc.12112

von der Pütten, A.M., Krämer, N.C., Gratch, J., Kang, S-H. (2010).”It doesn’t matter what you are!” Explaining social effects of agents and avatars. *Computers in Human Behavior, 26,* 1641-1650. doi: 10.1016/j.chb.2010.06.012

Preschl, B., Marecker, A., & Wagner, B. (2011). The working alliance in a randomized controlled trial comparing online with face-to-face cognitive-behavioral therapy for depression. *BMC Psychiatry, 11*, Article ID 189. doi: 10.1186/1471-244X-11-189

Qu, C., Brinkman, W-P., Ling, Y., Wiggers, P., Heyenderickx, I. Conversations with a virtual human: synthetic emotions and human responses, *Computers in Human Behavior, 34,* 58-68. doi: 10.1016/j.chb.2014.01.033

Ramirez, A., Walther, J.B., Burgoon, J.K., Sunnafrank, M. (2002). Information-seeking strategies, uncertainty, and computer-mediated communication. *Human Communication Research, 28*(2), 213–228. doi: 10.1111/j.1468-2958.2002.tb00804.x

Reeves, B., & Nass, C. (1996). *The media equation: how people treat computers, television, and media like real people and places*. New York, NY: Cambridge University Press.

Rehm, I.C. Foenander, E., Wallace, K., Abbott, J.M., Kyrios, M., Thomas, N. (2016). What role can avatars play in e-mental health interventions? Exploring new models of client-therapist interaction. *Frontiers in Psychiatry, 7*(3), 186. doi: 10.3389/fpsyt.2016.00186

Richards, D., & Richardson, T. (2012). Computer-based psychological treatments for depression: a systematic review and meta-analysis. *Clinical Psychology Review, 32,* 329-342. doi: 10.1016/j.cpr.2012.02.004

Rubin, D.R. (1976). Inference and missing data, *Biometrika, 63*(3), 581-592. Retrieved from https://www.jstor.org/stable/2335739

Russell, J.A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*(1), 102-141. doi: 10.1037/0033-2909.115.1.102

Schafer, J.L. (1999). Multipel imputation: a primer. *Statistical Methods in Medicine Research, 8,* 3-15. doi: 10.1177/096228029900800102

Schafer, J.L., & Graham, J.W. (2002). Missing data: our view of the state of the art. *Psychological Methods, 7*(2), 147-177. doi: 10.1037//1082-989X.7.2.147

Sommers-Flanagan, J. & Shaw, S. (2017). Clinical interviews. In A. Wenzel (Ed.), *The sage encyclopedia of abnormal and clinical psychology* (Vol. 1, pp. 683-686). Thousand Oaks,, CA: SAGE Publications, Inc. doi: 10.4135/9781483365817.n262

Stöckl., S., Schulte-Mecklenbeck, M., Borer, S., & Samsom, A.C. (2018). Facial expression analysis with AFFDEX and FACET: a validation study. *Behavior Research Methods, 50*(4), 1446-1460. doi: 10.3758/s13428-017-0996-1

Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics (2nd ed.)*. New York: Harper and Row.

Titov, N., Andrews, G., Choi, I., Schwencke, G., Mahoney, A. (2008). Shyness 3: Randomized controlled trial of guided versus unguided internet-based CBT for social phobia. *Australian & New Zealand Journal of Psychiatry*, *42*(12), 1030-1040. doi: 10.1080/00048670802512107

Titov, N., Andrews, G., Choi, I., Schwencke, G., Johnston, L. (2009a). Randomized controlled trial of web-based treatment of social phobia without clinician guidance. *Australian & New Zealand Journal of Psychiatry*, *43*, 913-919. doi: 10.1080/00048670903179160

Vedaa, Ø., Hagatun, S., Kallestad, H., Pallesen, S., Smith, O.R.F. (2019). Long-term effects of an unguided online cognitive behavioral therapy for chronic insomnia. *Journal of Clinical Sleep Medicine, 15*(1), 101-110. doi: 10.5664/jcsm.7580

Västfjäll, D., & Gärling, T. (2007). Validation of a Swedish short self-report measure of core affect. *Scandinavian Journal of Psychology, 48*(3)*,* 233-238. doi: 10.1111/j.1467-9450.2007.00595.x

Västfjäll, D., Friman, M., & Kleiner, M. (2002). The measurement of core affect: a Swedish self-report measure derived from the affect circumplex. *Scandinavian Journal of Psychology, 43*(1), 19-31. doi: 10.1111/1467-9450.00265

Wagner, B., Brand, J., Schulz, W., Knaevelsrud, C. (2012). Online working alliance predicts treatment outcome for posttraumatic stress symptoms in arab war-traumatized patients. *Depression and Anxiety, 29*(7), 646-651. doi: 10.1002/da.21962

Walther, J.B. (1992). Interpersonal effects in computer-mediated interaction: a relational perspective, *Communication Research, 19,* 52–90. doi: 10.1177/009365092019001003

Watson, J.C., Steckley, P.L., & McMullen, E.J. (2014). The role of empathy in promoting change. *Psychotherapy Research, 24*(3), 286-298. doi: 10.1080/10503307.2013.802823

World Health Organization. (2017). *Depression and other common mental health disorders: Global health estimates.* Retrieved from https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf

Appendix
Interview questions (in Swedish)

Main theme	*Interview questions and lines*
Introduction	Välkommen!
	Den här studien går ut på att undersöka olika psykologiska intervjuer online. Jag kommer att ställa frågor till dig och du ska svara på dem som du förstår dem. Det finns inga rätt eller fel svar på frågorna och du får själv bedöma hur mycket du vill svara. Ofta kommer vi vidare till följande fråga men ibland kommer jag att be dig berätta mera. Kom också ihåg att du får avbryta studien när som helst utan orsak, och du behöver heller inte förklara varför du avbryter.
	Har du förstått?
	Okej, bra. Är du ännu villig att fortsätta?
	Okej, vi kommer att fortsätta med intervjun. Först kommer jag att ställa frågor om dig och om din vardag.
Main interview	Berätta lite om dig själv.
	Hur ser en typisk dag ut för dig?
	Hur mycket tid spenderar på studierna eller arbetet?
	Hur mycket tid spenderar du med din partner och dina vänner?
	Vad brukar du göra tillsammans med din partner och dina vänner?
	Vad tycker du om att göra när du är ensam?
	Åsnebrygga: Vi har nu pratat lite allmänt om vad som händer i ditt liv just nu. Vi ska gå vidare och prata lite om hur du mår i allmänhet.
	Hur skulle du beskriva ditt allmänna mående?
	Hur är din fysiska hälsa (t.ex. aptit, sömn)?
	Hur skulle du beskriva din psykiska hälsa (t.ex. stress, humör)?
	Hur reagerar du på krav som ställs på dig i vardagen, kraven kan komma från andra eller dig själv?
	Vad brukar du göra för att återhämta dig från något som varit stressigt?
	Berätta lite om de viktigaste relationerna i ditt liv.
	Har du för tillfället några bekymmer i dina relationer?
	Hur nöjd känner du dig med ditt liv som det ser ut just nu?
Outro	Intervjun är nu slut. Assistenten kommer nu efteråt att diskutera lite med dig och det finns några blanketter för dig att fylla i ännu. Tack för att du deltog. Hejdå!

PRESSMEDDELANDE

En videobaserad interaktion väcker färre glada ansiktsuttryck och upplevs som mindre empatisk jämfört med en liveinteraktion

Pro gradu-avhandling i psykologi
Fakulteten för humaniora, psykologi och teologi vid Åbo Akademi

En pro gradu-avhandling i psykologi vid Åbo Akademi har undersökt effekten av ökad automatiseringsgrad av närvaro (i.e., mindre nyanserade sociala signaler) på emotionellt engagemang och upplevd empati i en intervju online. Syftet var att se om det finns skillnader mellan en liveinteraktion (t.ex. ett skypesamtal) och en videobaserad interaktion (t.ex. inspelade videoklipp som styrs av en människa), samt om det fanns skillnader mellan en skriptad liveinteraktion och en videobaserad interaktion. I studien jämförde man en semiskriptad liveintervju, en skriptad liveintervju och en inspelad videointervju och fann att deltagare oftare uttryckte glädje och tillskrev intervjuaren mer empatisk förmåga i liveintervjuerna än i den inspelade videointervjun. Man fann inga skillnader i upplevt emotionellt engagemang mellan grupperna. Resultaten tyder på att man uppfattar en videobaserad interaktion som mindre empatisk och uttrycker mindre glädje i interaktionen jämfört med en liveinteraktion, men att ökad automatiseringsgrad av närvaro inte påverkar upplevt emotionellt engagemang.

Resultaten kan bero på att den inspelade videointervjun upplevts som mindre realistisk än liveintervjuerna, men mera forskning behövs för att klargöra resultatens giltighet och om automatiserade interaktioner kan vara effektiva.

Studien hade sammanlagt 75 deltagare. Upplevd empati och emotionellt engagemang mättes med självskattningsformulär. Emotionella reaktioner analyserades av ett mjukvaruprogram som mätte deltagarnas ansiktsuttryck.

Ytterligare information fås av:
Anna-Karin Noromies
Magisterstuderande i psykologi
Psykologi/Åbo Akademi
0505670563
anoromie@abo.fi