



Reporting Mental Health Symptoms: Breaking Down Barriers to Care with Virtual Human Interviewers

Gale M. Lucas^{1*}, Albert Rizzo¹, Jonathan Gratch¹, Stefan Scherer¹, Giota Stratou¹, Jill Boberg¹ and Louis-Philippe Morency²

¹Institute for Creative Technologies, University of Southern California, Los Angeles, CA, United States, ²School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, United States

OPEN ACCESS

Edited by:

Mel Slater,
University of Barcelona, Spain

Reviewed by:

Lucia Valmaggia,
Institute of Psychiatry, Psychology &
Neuroscience (IoPPN),
United Kingdom
Xueni Pan,
Goldsmiths, University of London,
United Kingdom

*Correspondence:

Gale M. Lucas
lucas@ict.usc.edu

Specialty section:

This article was submitted
to Virtual Environments,
a section of the journal
Frontiers in Robotics and AI

Received: 26 July 2017

Accepted: 22 September 2017

Published: 12 October 2017

Citation:

Lucas GM, Rizzo A, Gratch J,
Scherer S, Stratou G, Boberg J and
Morency L-P (2017) Reporting
Mental Health Symptoms:
Breaking Down Barriers to Care
with Virtual Human Interviewers.
Front. Robot. AI 4:51.
doi: 10.3389/frobt.2017.00051

A common barrier to healthcare for psychiatric conditions is the stigma associated with these disorders. Perceived stigma prevents many from reporting their symptoms. Stigma is a particularly pervasive problem among military service members, preventing them from reporting symptoms of combat-related conditions like posttraumatic stress disorder (PTSD). However, research shows (increased reporting by service members when anonymous assessments are used. For example, service members report more symptoms of PTSD when they *anonymously* answer the Post-Deployment Health Assessment (PDHA) symptom checklist compared to the *official* PDHA, which is identifiable and linked to their military records. To investigate the factors that influence reporting of psychological symptoms by service members, we used a transformative technology: automated virtual humans that interview people about their symptoms. Such virtual human interviewers allow simultaneous use of two techniques for eliciting disclosure that would otherwise be incompatible; they afford *anonymity* while also building *rapport*. We examined whether virtual human interviewers could increase disclosure of mental health symptoms among active-duty service members that just returned from a year-long deployment in Afghanistan. Service members reported more symptoms during a conversation with a virtual human interviewer than on the official PDHA. They also reported more to a virtual human interviewer than on an *anonymized* PDHA. A second, larger sample of active-duty and former service members found a similar effect that approached statistical significance. Because respondents in both studies shared more with virtual human interviewers than an anonymized PDHA—even though both conditions control for stigma and ramifications for service members' military records—*virtual human interviewers that build rapport* may provide a superior option to encourage reporting.

Keywords: virtual humans, assessment, disclosure, psychological symptoms, anonymity

INTRODUCTION

People are reluctant to disclose information that could be potentially stigmatizing. One area where this failure to disclose honest information has particularly large consequences is mental health. Due to the stigma associated with mental health problems (Link et al., 1991, 2001), people are reluctant to report symptoms of such disorders. The consequences are significant—mental health problems exact

a significant toll on society (World Health Organization, 2004; Insel, 2008; National Institute of Mental Health, 2010; World Economic Forum, 2011).

The majority of individuals who seek mental health services report facing stigma and discrimination (Thornicroft et al., 2009). Accordingly, stigma and discrimination acts as a significant barrier to care and honest reporting of symptoms, which individuals must overcome. They may try to deal with stigma using coping methods that are more or less effective (Isaksson et al., 2017); however, if they cannot successfully cope, stigma and the resultant unwillingness to report symptoms end up preventing people from accessing or receiving treatment, leaving the disorder unresolved. These barriers to care pose a large problem for society since mental health problems are costly, both in terms of money and social capital (Insel, 2008) and unresolved mental health problems continue to accrue increasing costs (World Health Organization, 2004; Insel, 2008; National Institute of Mental Health, 2010; World Economic Forum, 2011).

In the current work, we explore this problem among military service members, given that failure to disclose symptoms is often cited as a barrier to care in the military [Institute of Medicine (IOM), 2014; Rizzo and Shilling, in press]. Service members are reluctant to report symptoms of combat-related conditions like posttraumatic stress disorder (PTSD), which is typified by persistent mental, behavioral, and emotional symptoms as a result of exposure to physical or psychological trauma. Not only are service members more likely to have PTSD than civilians (Vincenzes, 2013; Schreiber and McEnany, 2015) but also as a result of the perceived stigma surrounding the condition (Hoge et al., 2004, 2006), they are particularly reluctant to report symptoms (Olson et al., 2004; Appenzeller et al., 2007; Warner et al., 2007, 2008, 2011; McLay et al., 2008; Fear et al., 2010; Thomas et al., 2010). The reluctance of service members in the United States Military to report PTSD symptoms is especially intensified when they are screened for mental health symptoms using the official administration of the Post-Deployment Health Assessment (PDHA; Hyams et al., 2002; Wright et al., 2005) since this information becomes documented in their military health records. Indeed, there are pragmatic military career implications (such as the perception of possible future restrictions from certain job placements and from obtaining future security clearances) for having been screened positive for mental health conditions.

To address this reluctance to disclose PTSD symptoms on the PDHA, we examine whether a new technology, namely virtual human-interviewers, can be used to increase willingness of service members to report PTSD symptoms compared to the PDHA. Virtual humans are digital representations of humans that can portray human-like characteristics and abilities and can be used to interview people in a natural way using conversational speech (see **Figure 1**). We first describe empirical work that provides a theoretical basis for how virtual human interviewers might increase the willingness of service members to report PTSD symptoms, followed by the research questions addressed in the current work. We then describe and discuss results from two studies that examine the effectiveness of virtual human interviewers designed to foster service member reporting of mental

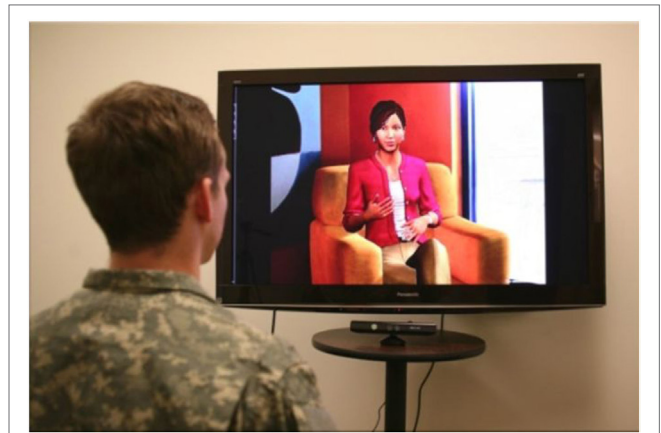


FIGURE 1 | Interview with our virtual human interviewer.

health symptoms that might otherwise be withheld when using traditional self-report checklists (such as the PDHA).

Related Work

Anonymity is theorized to support the potential effectiveness of virtual human interviewers for increasing reports of health-related symptoms. Previous research suggests that anonymized forms of assessment increase reporting. For example, respondents reveal more honest information during computerized self-assessments (Greist et al., 1973; Beckenbach, 1995; Weisband and Kiesler, 1996; van der Heijden et al., 2000; Joinson, 2001), and they appear to do so because they perceive these assessments to be more anonymous than non-computerized human interviewing methods (Sebestik et al., 1988; Thornberry et al., 1990; Baker, 1992; Beckenbach, 1995; Joinson, 2001). Although anonymized assessments can improve honest reporting for even mundane private information (Beckenbach, 1995; Joinson, 2001), these effects are especially strong when the information is illegal, unethical, or culturally stigmatized (Weisband and Kiesler, 1996; van der Heijden et al., 2000). As many behaviors that are harmful to mental and physical health fall into this category (e.g., drug use, unsafe sex, suicide attempts), anonymized forms of assessment can be especially important in healthcare assessment. For example, when asked to disclose information about suicidal thoughts using a computer-administered assessment, participants not only felt more positive about the assessment compared to traditional methods, but also gave more honest answers (Greist et al., 1973).

Relevant to the focus of the current work, anonymity has been shown to increase reporting disclosure of PTSD symptoms among service members (Olson et al., 2004; McLay et al., 2008; Warner et al., 2011). One study indicated that following a combat deployment, the sub-sample of service members who *anonymously* answered the routine PDHA symptom checklist reported twofold to fourfold higher mental health symptoms and a higher interest in receiving care compared to the overall results derived from the standard administration of the PDHA, which is identifiable and linked to service members' military records (Warner et al., 2011).

Initial research on the use of virtual humans to conduct clinical interviews suggests that interviewees are indeed more open to virtual human interviewers than their human counterparts (Slack and Van Cura, 1968; Lucas et al., 2014; Pickard et al., 2016). Because a conversation with a virtual human interviewer may be viewed as more anonymous, users may be more comfortable disclosing about highly sensitive topics and on questions that could lead them to admit something stigmatized or otherwise negative. For example, during a clinical interview with a virtual human interviewer, participants disclose more personal details when they are told that the virtual human is autonomous than when they are told that the virtual human is operated by a person in another room (Lucas et al., 2014). Pickard et al. (2016) reported that individuals are more comfortable disclosing to an automated virtual human interviewer than its human counterpart.

While research has yet to establish that virtual human interviewers can increase reporting of PTSD symptoms specifically among service members, some research has considered the potential benefits of using virtual human interviewers and related technology for service members (Lewandowski et al., 2011; Rizzo et al., 2011; Serowik et al., 2014; Bhalla et al., 2016). Rizzo et al. (2011) developed a virtual human to interview service members about their PTSD symptoms. Advances in automation now allow virtual human interviewers to have more interactive conversations with users, in which questions about the PTSD symptoms can be embedded. Having such an interactive conversation is critical because, while anonymity is beneficial, building rapport with respondents can also increase reporting (Burgoon et al., 2016).

Indeed, a second theoretical basis behind the potential effectiveness of virtual human interviewers for increasing report of symptoms is rapport. Psychological theories of rapport (e.g., Tickle-Degnen and Rosenthal, 1990) have outlined verbal and non-verbal behaviors that help to build rapport; and subsequent research has shown that resultant rapport leads interlocutors to disclose more (Miller et al., 1983; Hall et al., 1996; Gratch et al., 2007, 2013; Burgoon et al., 2016). Differences in disclosure between assessment formats have also been found to be mediated by feelings of rapport; rapport leads individuals to disclose more personal information (Dijkstra, 1987; Gratch et al., 2007, 2013).

Because traditional computerized self-assessments and other anonymized forms lack any human element, these traditional assessments do not evoke the same feelings of rapport or social connection. Specifically, when there is not a human or human-like agent present in some way, shape, or form, people feel less socially connected during the assessment (Gratch et al., 2007, 2013).

Tickle-Degnen and Rosenthal (1990) suggest several features of “the human element” that are important in increasing rapport, including both verbal and non-verbal behavior. For example, listeners who are naturally more verbally receptive and attentive and who use more follow-up questions, produce greater disclosure from reticent interviewees (Miller et al., 1983). Beyond the words uttered, non-verbal behavior such as positive facial expressions, attentive eye gaze, welcoming gestures and open postures have been reported to influence feelings of rapport (Hall et al., 1996; Burgoon et al., 2016). These features may allow virtual human interviewers to more effectively build rapport, in contrast to

traditional computerized self-assessments and other anonymized forms. Indeed, research suggests that virtual human interviewers have the potential build rapport as well as—or even better than—human interviewers (e.g., DeVault et al., 2014).

Researchers have attempted to translate these psychological theories of rapport into computational systems and studies have indicated that it is possible to capture these behaviors in various automated systems ranging from machine learning-based prediction models (e.g., Morency et al., 2009; Huang et al., 2010) to “chatbots” (e.g., Kerlyl et al., 2007) and virtual humans (e.g., Cassell and Bickmore, 2002; Bickmore et al., 2005; Haylan, 2005; Cassell et al., 2007; Gratch et al., 2007, 2013; Matsuyama et al., 2016; Zhao et al., 2016). For example, some virtual humans have been designed to utilize verbal (e.g., words uttered, prosody, intonation, etc.) and non-verbal behavior (e.g., positive facial expressions, gaze, gestures, and posture) to build rapport (Cassell and Bickmore, 2002; Bickmore et al., 2005; Haylan, 2005; Cassell et al., 2007; Gratch et al., 2007, 2013; Matsuyama et al., 2016; Zhao et al., 2016). Research has also established that virtual humans that employ such rapport-building behaviors are able to induce disclosure (Gratch et al., 2007, 2013).

While rapport-building seems contrary to anonymity, the use of virtual human interviewers may provide a solution that allows for both anonymity as well as rapport-building. Some virtual humans can be used to interview people in a natural way (i.e., *via* conversational speech). Akin to the “Rapport Agents” described above, these virtual human interviewers have been designed to *build rapport with users specifically during interviews* (e.g., Gratch et al., 2013; Qu et al., 2014), including clinical interviews (e.g., Bickmore et al., 2005; DeVault et al., 2014; Lucas et al., 2014; Rizzo et al., 2016). Interspersed *appropriately* during an interview, the virtual human interviewers use verbal and non-verbal backchannels (e.g., utterances of agreement such as “mhm” or head nods) to build rapport with the interviewee. Indeed, virtual human interviewers that employ such backchannels when appropriate to the conversation create greater feelings of rapport than virtual human interviewers that employ them at random during the interview (e.g., Gratch et al., 2013; Qu et al., 2014). As with Rapport Agents, when virtual human interviewers use these rapport-building behaviors in this way, they are able to prompt disclosure from interviewees (Gratch et al., 2013).

The Current Research

Given that the experience of stigma can limit the reporting of PTSD symptoms, many service members with the disorder are not identified and do not have the opportunity to benefit from the evidence-based treatments that currently exist. By using a virtual human interviewer to increase self-disclosure of more accurate information, service members having such difficulties could be better encouraged to access potentially beneficial mental health care options. Although the prior research is suggestive, it has not sufficiently established that virtual human interviewers can be used to increase service members’ willingness to endorse the presence of PTSD symptoms compared to self-report on the PDHA checklist items. Thus, the current study tested whether virtual human interviewers can encourage reporting of PTSD symptoms compared to the gold-standard PHDA. Accordingly,

we hypothesize that service members will be more willing to disclose PTSD symptoms to a virtual human interviewer than on the official PDHA (H1). The study also examines the role of rapport *in addition to* anonymity for increasing disclosure. While Warner et al. (2011) demonstrated that service members who answered the PDHA symptom checklist anonymously were more willing to report mental health symptoms compared to the official PDHA, virtual human interviewers with the added benefit of rapport-building, may have the capability to evoke higher levels of disclosure of symptoms. If rapport has an impact on self-disclosure in this context above and beyond anonymity, service members will be more willing to report symptoms to a virtual human interviewer than on an anonymized version of the PDHA (even though they are both equally anonymous). Thus, we hypothesized that service members would be more willing to report PTSD symptoms to a virtual human interviewer than on an anonymized version of the PDHA (H2).

In order to address these research questions, we conducted two studies to test whether service members (Studies 1 and 2) and Veterans (in Study 2) would be more willing to report PTSD symptoms when asked by a virtual human interviewer than when asked to report on the PDHA (either official or anonymized).

STUDY 1—MATERIALS AND METHODS

Study 1—Participants

In Study 1, 29 (2 females) active-duty Colorado National Guard service members volunteered to participate in the study during 2013. None of the service members in the unit declined to participate. After returning from a year-long deployment to Afghanistan, they completed the measures described below. The sample was diverse regarding age ($M = 41.46$, Range = 26–56) and previous number of combat deployments ($M = 2.00$, Range = 1–7). Due to technical failures five participants (all male) were excluded from the analysis reported below.

Study 1—Design and Procedure

This study compared reporting of PTSD symptoms in three formats: (1) standard administration of the PDHA upon return from deployment; (2) an anonymized version of the PDHA; and (3) PDHA questions on PTSD symptoms asked by a virtual human interviewer that were embedded in a longer set of general interview questions. All participants completed the official PDHA within 2 days of the other two assessments (either before or after these other two assessments) and signed releases to allow the research team to access their official PDHA responses gathered at post-deployment processing. Three questions on the PDHA assess whether the service member is experiencing the three core Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) diagnostic symptoms for PTSD (intrusive recollections; avoidance/numbing; hyperarousal).

On the official PDHA, participants were asked “*Have you ever had any experience that was so frightening, horrible, or upsetting that, in the past month, you:*

- (A) *have had nightmares about it or thought about it when you did not want to?* (intrusive recollection)
- (B) *tried hard not to think about it or went out of your way to avoid situations that remind you of it?* (avoidance/numbing)
- (C) *were constantly on guard, watchful, or easily startled?* (hyperarousal).

Participants selected “yes” or “no” on each of these three items, and their answers were submitted to the US Military as part of their official military health record. At the same time, we were granted access to these official PDHA responses for our study sample.

Next, participants arrived at the study site, gave consent, completed a demographic questionnaire, and rated their mood using items “I am happy” and “I worry too much” on 4-point scales from *almost never* to *almost always*. They then were escorted to a private room and completed the anonymized PDHA PTSD questions on a computer, selecting Yes or No responses to each item. The participants were verbally assured that their responses were confidential as they would be deidentified using a participant number code.

Participants were then engaged by a virtual human interviewer who conducted a semistructured screening interview *via* spoken language. Participants were still alone in the private room and were told they would not be observed by anyone during the interview and that the video recordings of their interview session would not be released to anyone outside the research team. The full interview was structured around a series of agent-initiated questions organized into three phases: Phase 1 was a rapport-building phase where the virtual human interviewer asked participants general introductory questions (e.g., “Where are you from originally?”); Phase 2 was the clinical phase where the virtual human interviewer asked a series of questions about symptoms (e.g., “How easy is it for you to get a good night’s sleep?”), which included the naturally embedded PDHA questions; Phase 3 was the ending section of the interview where the virtual human interviewer asks questions designed to return the patient to a more positive mood (e.g., “What are you most proud of?”). Across the session, the virtual human interviewer built rapport using follow-up questions (e.g., “Can you tell me more about that?”), empathetic feedback (e.g., “I’m sorry to hear that”), and non-verbal behaviors (e.g., nods, expressions).

The PDHA questions that were asked by the virtual human interviewer were slightly re-worded in order to embed them in the interview. In place of the three PDHA questions listed above, the virtual human interviewer asked participants these revised versions:

- (A) “*Can you tell me about any bad dreams you’ve had about your experiences, or times when thoughts or memories just keep going through your head when you wish they wouldn’t?*” (intrusive recollection)
- (B) “*Can you tell me about any times you found yourself actively trying to avoid thoughts or situations that remind you of past events?*” (avoidance/numbing)
- (C) “*Can you tell me about any times recently when you felt jumpy or easily startled?*” (hyperarousal).

Participants’ answers to the three PDHA questions during the interview were recorded and later coded by two blind coders as to

whether the participant had this experience in the last month or not. While more nuanced than checking “yes” or “no” on PDHA, our coders dichotomized open-ended responses to parallel the PDHA. For example, one response to the intrusive recollections question that was coded as “no” was “Um... haven’t had any, you know, dreams or nightmares. No.” A response to the avoidance question that was coded as “yes” stated: “Um... I try to leave early. I try to leave a situation. I try not to talk to those people. Um... That’s the only time I really avoid a situation is avoiding those people.” Coders had 100% agreement, and codes served as “yes” or “no” answers.

STUDY 1—RESULTS

In Study 1, three versions of the PDHA (official PDHA, Anonymized PDHA, and virtual human interviewer) were administered to participants to determine whether manner of administration produced differing responses. Scores were created for each version of the PDHA by counting the number of “yes” answers to the three questions that assess the core DSM-IV-TR diagnostic symptoms for PTSD (intrusive recollection, avoidance/numbing, hyperarousal). To compare responding, we conducted a repeated-measures ANOVA using 24 participants from a sample of active-duty Colorado National Guard who completed all three measures. There was a significant effect of assessment type, $F(2,23) = 4.29, p = 0.02$ (Figure 2). Within-subject contrasts revealed that participants reported more symptoms of PTSD (responded “yes” on more questions) when asked by the virtual human interviewer ($M = 0.79, SE = 0.23$) than when reporting on the official PDHA ($M = 0.25, SE = 0.15$), $F(1,23) = 7.38, p = 0.01$, or even when reporting on the anonymized version of the PDHA ($M = 0.33, SE = 0.16$), $F(1,23) = 4.84, p = 0.04$. The difference between official and anonymized versions of the PDHA was not significant [$F(1,23) = 0.19, p = 0.66$].

Study 1 provided an initial test of our research hypotheses with results suggesting that service members are more willing to report PTSD symptoms to a virtual human interviewer than

on the official PDHA (H1). The results also indicate that service members are more willing to report PTSD symptoms to a virtual human interviewer than on an anonymized version of the PDHA (Q2). Indeed, because respondents in this study shared more with virtual human interviewers than an anonymized PDHA—even though both conditions control for stigma and ramifications for service members’ military records—*virtual human interviewers that build rapport* may provide a superior option for encouraging endorsement of these symptoms. This finding has important implications, suggesting that virtual human interviewers may help service members “open up” and report their psychological symptoms through rapport building. We then conducted a second study to replicate (and extend) this result in a larger, more diverse sample including both active-duty service members and retired military veterans. In this second study, we also ruled out the confound introduced by the wording differences between the virtual human interviewer’s questions and the questions listed on the PDHA. In Study 2, the questions asked by the virtual human interviewer were worded identically to the questions on the anonymized PDHA.

STUDY 2—MATERIALS AND METHODS

Study 2—Participants

In Study 2, 132 (16 female) active duty service members and veterans were recruited (e.g., through *Craigslist*), and paid \$30 for their participation during 2014 and 2015. Only individuals who were enrolled as a part of the US military, either currently or in the past, were invited to participate. As in Study 1, this sample was diverse regarding age ($M = 44.12, \text{Range} = 18\text{--}77$), but information regarding number of deployments was not taken for this sample.

Study 2—Design and Procedure

Participants completed the same procedures as Study 1 with a few exceptions. First, since this sample included veteran participants who had not just returned from a deployment, we did not collect the official PDHA for this study. Thus, we only compared responses to the anonymized PDHA with the same questions asked by the virtual human interviewer.

Second, after giving consent and completing demographic questions, participants also completed additional screening measures including the PTSD Checklist (PCL; Blanchard et al., 1996). The PCL is a self-report measure that evaluates PTSD using a 5-point Likert scale. It is based on the DSM-IV-TR. Scores range from 17 to 85 and symptom severity is reflected in the size of the score, with larger scores indicating greater severity of PTSD symptoms. The PCL is commonly used in clinical practice and in research studies on PTSD. Participants also completed additional individual difference questionnaires that were not relevant to the current research questions, but described elsewhere (DeVault et al., 2014; Gratch et al., 2014).

Finally, and most importantly, Study 2 rules out the confound of question wording. While in Study 1 our anonymized version of the PDHA used the question wording from the official PHDA, in Study 2 our anonymized version of the PDHA used the exact

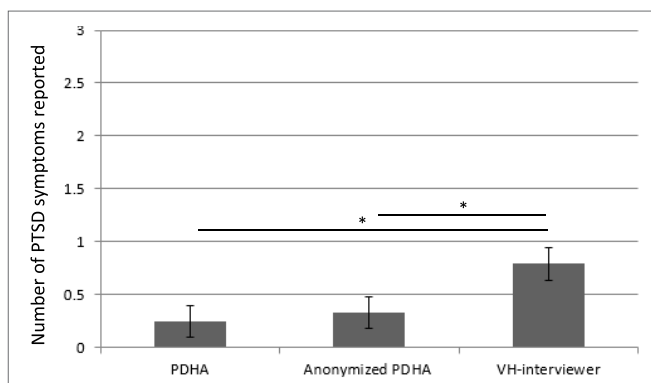


FIGURE 2 | Number of posttraumatic stress disorder (PTSD) symptoms reported out of three questions representing Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) categories. Participants reported fewer symptoms (1) on the official Post-Deployment Health Assessment (PDHA) collected by the Joint Forces and (2) on an anonymized version of the PDHA, than (3) to a virtual human interviewer during a postdeployment interview in Study 1. * $p < 0.05$.

wording employed by the virtual human interviewer (see above). Therefore, in this study, any differences observed between answers on the anonymized version of the PDHA and the interview led by a virtual human could not be due to question wording. Coders again dichotomized each response as “yes” or “no” for the PTSD symptom, and had 100% agreement.

STUDY 2—RESULTS

A repeated-measures *t*-test among participants who successfully completed both assessments ($n = 126$) revealed an effect of assessment type that approached statistical significance [$t(125) = 1.76$, $p = 0.08$]; participants reported more PTSD symptoms when asked by the virtual human interviewer ($M = 1.21$, $SE = 0.10$) than on an anonymized version of the PDHA ($M = 1.05$, $SE = 0.11$; **Figure 3A**). There is a significant interaction with PCL score [$F(1,124) = 4.38$, $p = 0.04$] such that among those with subtler subthreshold PTSD symptoms (below median on the PTSD Checklist; PCL; Blanchard et al., 1996), the effect is significant [$M = 0.53$, $SE = 0.10$ vs. $M = 0.17$, $SE = 0.07$; $t(63) = 3.77$, $p < 0.001$; **Figure 3B**]. However, there is no significant difference among those already reporting higher symptoms on the PCL [$M = 1.92$, $SE = 0.12$ vs. $M = 1.95$, $SE = 0.14$; $t(61) = -0.20$,

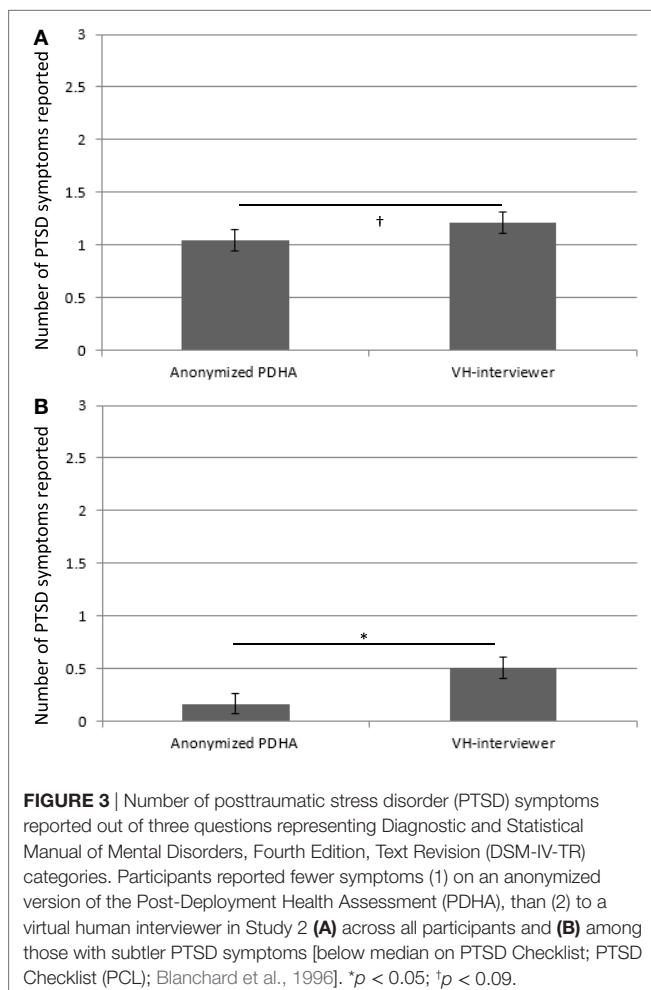
$p = 0.84$]. Likewise, entering PCL score as a covariate rendered the aforementioned effect of assessment type on number of reported symptoms significant [virtual human interviewer $M = 1.21$, $SE = 0.08$ vs. anonymized PDHA $M = 1.05$, $SE = 0.08$; $F(1,124) = 5.78$, $p = 0.02$]. Finally, while an ANOVA revealed a significant between-subjects main effect of active duty status on number of reported symptoms such that active duty subjects were overall less willing to report symptoms ($M = 0.16$, $SE = 0.25$) than veterans [$M = 1.27$, $SE = 0.10$; $F(1,124) = 17.32$, $p < 0.001$], there was no interaction between assessment type and active duty status [$F(1,124) = 0.34$, $p = 0.56$].

Like Study 1, Study 2 demonstrates that service members are more willing to report PTSD symptoms to a virtual human interviewer than on an anonymized version of the PDHA (Q2). Indeed, even though both conditions control for stigma and ramifications for service members' military records, participants are more willing to report PTSD symptoms to virtual human interviewer than an anonymous version of the PDHA.

DISCUSSION

Across both studies, participants reported more PTSD symptoms when asked by a virtual human interviewer. Study 1 showed the effectiveness of virtual human interviewers in a sample of active duty service members reporting symptoms of mental distress. Supporting H1, service members reported more symptoms during a conversation with a virtual human interviewer than on the official PDHA. Our analysis of the small sample in Study 1 did not reveal differences between official and anonymized versions of the PDHA as was reported in Warner et al. (2011). However, in Warner et al., within-group results were not assessed and instead mean group differences between those who volunteered to fill out the anonymous version were compared with the mean of the larger official PDHA sample (1,712 out of 3,502). Service members in Study 1 also reported more PTSD symptoms to a virtual human interviewer than on an *anonymized* PDHA. In Study 2, we found a similar effect that approached statistical significance using a larger sample of active-duty service members and veterans. As in Study 1, participants in this study tended to report more symptoms when asked by a virtual human interviewer than on an anonymized PDHA. Thus, both reported studies support H2. Furthermore, the second study suggests that individuals falling under the radar in traditional assessments and scoring low on questionnaires like the PCL (e.g., possibly due to impression management, fear of stigmatization) could be detected by virtual human interviewers. Indeed, in this second study (where the sample has a broader range of distress), without taking into account PCL, the effect of assessment type on reporting of PTSD symptoms only approached statistical significance.

Although we showed that virtual human interviewers can increase service members' disclosure of mental health symptoms, further research is required to rule out alternative explanations concerning the mechanism behind this disclosure. For example, the open-ended nature of the questions asked by the virtual human interviewer could have contributed to encouraging service members to disclose. To see the extent to which this factor



contributes, future research could—for example—compare an open-ended paper-and-pencil version of the PDHA questions to the official forced-choice version in the absence of rapport building. Likewise, in both studies, all participants completed the anonymized PDHA before the interview with the virtual human, leaving order as another possible alternative explanation. However, this is unlikely to explain our results because, in Study 1, some participants completed the official PDHA before the anonymized PDHA and the virtual human interview, whereas others completed the official PDHA *after* these other two assessments. Although we do not have access to the dates when specific service members in our study completed the official PDHA to further test this, if order made a significant contribution, we would not have found the strongest effect of assessment (1) in this study and (2) when comparing to this (official) assessment, which was completed last for some participants. While this may help to rule out order as an alternative explanation for the difference between the virtual human interviewer and the official PDHA, it does not preclude the possibility that an order effect contributed to the difference between the virtual human interviewer and the anonymized PDHA.

In line with previous studies (Slack and Van Cura, 1968; Lucas et al., 2014; Pickard et al., 2016), these results support the view that virtual human interviewers provide a safe, reduced-stigma context where users may reveal more honest information. However, our results also go beyond prior work in that the current study focused specifically on service members and veterans, rather than a general civilian population. Also, where other clinical interviews led by virtual humans are more general, the clinical interview in this work assessed responses to specific questions about symptoms of PTSD. Thus, the results of this study add to previous work on use of such technologies for service members by demonstrating that virtual human interviewers may have a role to play in enhancing military mental health assessment by encouraging service members to report more PTSD symptoms than the gold-standard PDHA.

Moreover, beyond effects of anonymity found previously (e.g., Warner et al., 2011), virtual human interviewers may help soldiers “open up” and report their psychological symptoms through rapport building. Given that service members were more willing to report symptoms to a virtual human interviewer than on an anonymized version of the PDHA—even though these assessments were equally anonymous, this work establishes the idea that rapport has an impact on self-disclosure *above and beyond anonymity*. Pragmatically, this finding makes the case for taking advantage of the value that rapport-building holds for honest reporting rather than just relying on anonymity. For example, just having an anonymous online form appears not to be a sufficient “technological leap” to maximize self-disclosure. Honest reporting of such symptoms can better inform accurate diagnosis and help service members and civilians to break down barriers to care and receive evidence-based interventions that could mitigate the serious consequences of having a chronic untreated health condition. As such, the benefits of virtual human administrated mental health assessments could be substantial.

Finding that there is an impact of rapport for increasing disclosure *in addition to* anonymity has implications beyond

just reporting of psychological symptoms. Building upon the established effect of anonymity on disclosure (Sebestik et al., 1988; Thornberry et al., 1990; Baker, 1992; Beckenbach, 1995; Joinson, 2001; Warner et al., 2011), rapport-building could be beneficial for honest disclosure of any kind of sensitive information. As reviewed by Weisband and Kiesler (1996), anonymous assessments are especially helpful for eliciting information that is illegal (such as crimes like sexual assault) or is largely considered unethical or at least taboo (like risky sexual activity); our work implies that adding the second technique to elicit disclosure of rapport-building would further increase honest reporting of such information. Because virtual humans can build rapport while maintaining anonymity, they could be particularly useful for encouraging these kinds of disclosures.

Future work should investigate the impact of virtual human interviewers on promoting honest disclosure in other such sensitive clinical domains (Rizzo and Koenig, *in press*). Additionally, virtual human interviewers could be considered as an assessment strategy in other areas (e.g., financial planning) where people may perceive at least *some* stigma, and therefore may be tempted to under-report certain values (such as debt) even though honest information is essential for practitioners to give clients sound advice. Virtual human interviewers might also be useful for gaining honest information that—while not particularly stigmatizing—is still uncomfortable to disclose. For example, in organizational contexts, virtual humans could be helpful in eliciting honest performance evaluations.

ETHICS STATEMENT

The study was approved by University of Southern California's Institutional Review Board. Prior to participating, every participant consented by signing a written informed consent. No vulnerable populations were involved.

AUTHOR CONTRIBUTIONS

GL made substantial contributions to the conception and design of the work as well as the analysis or interpretation of data. AR and JG made substantial contributions to the conception and design of the work as well as interpretation of data. SS and GS made substantial contributions to the conception of the work as well as analysis and interpretation of data. JB made substantial contributions to the acquisition, analysis, or interpretation of data. L-PM made substantial contributions to the conception and design of the work.

FUNDING

This work was supported by DARPA under contract W911NF-04-D-0005 and the US Army. Any opinion, content or information presented does not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

REFERENCES

- Appenzeller, G. N., Warner, C. H., and Grieger, T. (2007). Postdeployment health reassessment: a sustainable method for brigade combat teams. *Mil. Med.* 172, 1017–1023. doi:10.7205/MILMED.172.10.1017
- Baker, R. P. (1992). New technology in survey research: computer-assisted personal interviewing (CAPI). *Soc. Sci. Comput. Rev.* 10, 145–157. doi:10.1177/089443939201000202
- Beckenbach, A. (1995). Computer-assisted questioning: the new survey methods in the perception of the respondents. *Bull. Method Soc.* 48, 82–100. doi:10.1177/075910639504800111
- Bhalla, A., Durham, R. L., Al-Tabaa, N., and Yeager, C. (2016). The development and initial psychometric validation of the eHealth readiness scale. *Comput. Human Behav.* 65, 460–467. doi:10.1016/j.chb.2016.09.015
- Bickmore, T., Gruber, A., and Picard, R. (2005). Establishing the computer–patient working alliance in automated health behavior change interventions. *Patient Educ. Couns.* 59, 21–30. doi:10.1016/j.pec.2004.09.008
- Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., and Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behav. Res. Ther.* 34, 669–673. doi:10.1016/0005-7967(96)00033-2
- Burgoon, J. K., Guerrero, L. K., and Floyd, K. (2016). *Nonverbal Communication*. London: Routledge.
- Cassell, J., and Bickmore, T. (2002). Negotiated collusion: modeling social language and its relationship effects in intelligent agents. *User Model. Adap. Interface.* 12, 1–44. doi:10.1023/A:1013337427135
- Cassell, J., Gill, A., and Tepper, P. (2007). “Conversational coordination and rapport,” in *Proceedings of Workshop on Embodied Language Processing at ACL 2007* (Prague, CZ).
- DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., et al. (2014). “SimSensei Kiosk: a virtual human interviewer for healthcare decision support,” in *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems* (Paris: International Foundation for Autonomous Agents and Multiagent Systems), 1061–1068.
- Dijkstra, W. (1987). Interviewing style and respondent behavior an experimental study of the survey-interview. *Soc. Method Res.* 16, 309–334. doi:10.1177/0049124187016002006
- Fear, N. T., Jones, M., Murphy, D., Hull, L., Iversen, A. C., Coker, B., et al. (2010). What are the consequences of deployment to Iraq and Afghanistan on the mental health of the UK armed forces? A cohort study. *Lancet* 375, 1783–1797. doi:10.1016/S0140-6736(10)60672-1
- Gratch, J., Artstein, R., Lucas, G. M., Stratou, G., Scherer, S., Nazarian, A., et al. (2014). “The distress analysis interview corpus of human and computer interviews,” in *Proceedings of the 9th Language Resources Evaluation Conference*, Reykjavik, 3123–3128.
- Gratch, J., Kang, S. H., and Wang, N. (2013). “Using social agents to explore theories of rapport and emotional resonance,” in *Emotion in Nature and Artifact*, eds J. Gratch and S. Marsella (Oxford: Oxford University Press), 181–197.
- Gratch, J., Wang, N., Gerten, J., Fast, E., and Duffy, R. (2007). “Creating rapport with virtual agents,” in *International Workshop on Intelligent Virtual Agents* (Berlin Heidelberg: Springer), 125–138.
- Greist, J. H., Laughren, T. P., Gustafson, D. H., Staus, F. F., Rowse, G. L., and Chiles, J. A. (1973). A computer interview for suicide-risk prediction. *Am. J. Psychiatry* 130, 1327–1332. doi:10.1176/ajp.130.12.1327
- Hall, J. A., Harrigan, J. A., and Rosenthal, R. (1996). Nonverbal behavior in clinician–patient interaction. *Appl. Prev. Psychol.* 4, 21–37. doi:10.1016/S0962-1849(05)80049-6
- Haylan, D. (2005). *Challenges Ahead. Head Movements and Other Social Acts in Conversation*. Hertfordshire, UK: AISB.
- Hoge, C. W., Auchterlonie, J. L., and Milliken, C. S. (2006). Mental health problems, use of mental health services, and attrition from military service after returning from deployment to Iraq or Afghanistan. *JAMA* 295, 1023–1032. doi:10.1001/jama.295.9.1023
- Hoge, C. W., Castro, C. A., Messer, S. C., McGurk, D., Cotting, D. I., and Koffman, R. L. (2004). Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *N. Engl. J. Med.* 351, 13–22. doi:10.1056/NEJMoa040603
- Huang, L., Morency, L. P., and Gratch, J. (2010). “Learning backchannel prediction model from parasocial consensus sampling: a subjective evaluation,” in *10th International Conference on Intelligent Virtual Agents* (Philadelphia, PA).
- Hyams, K. C., Riddle, J., Trump, D. H., and Wallace, M. R. (2002). Protecting the health of United States military forces in Afghanistan: applying lessons learned since the Gulf War. *Clin. Infect. Dis.* 34(Suppl. 5), S208–S214. doi:10.1086/340705
- Insel, T. R. (2008). Assessing the economic costs of serious mental illness. *Am. J. Psychiatry* 165, 663–665. doi:10.1176/appi.ajp.2008.08030366
- Institute of Medicine (IOM). (2014). *Treatment for Posttraumatic Stress Disorder in Military and Veteran Populations: Final Assessment*. Washington, DC: The National Academies Press.
- Isaksson, A., Corker, E., Cotney, J., Hamilton, S., Pinfold, V., Rose, D., et al. (2017). Coping with stigma and discrimination: evidence from mental health service users in England. *Epidemiol. Psychiatr. Sci.* 1–12. doi:10.1017/S204579601700021X
- Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: the role of self-awareness and visual anonymity. *Eur. J. Soc. Psychol.* 31, 177–192. doi:10.1002/ejsp.36
- Kerlyl, A., Hall, P., and Bull, S. (2007). “Bringing chatbots into education: towards natural language negotiation of open learner models,” in *Applications and Innovations in Intelligent Systems XIV*, ed. R. Ellis (London: Springer), 179–192.
- Lewandowski, J., Rosenberg, B. D., Parks, M. J., and Siegel, J. T. (2011). The effect of informal social support: Face-to-face versus computer-mediated communication. *Comput. Human Behav.* 27, 1806–1814. doi:10.1016/j.chb.2011.03.008
- Link, B. G., Mirotnik, J., and Cullen, F. T. (1991). The effectiveness of stigma coping orientations: can negative consequences of mental illness labeling be avoided? *J. Health Soc. Behav.* 32, 302–320. doi:10.2307/2136810
- Link, B. G., Struening, E. L., Neese-Todd, S., Asmussen, S., and Phelan, J. C. (2001). Stigma as a barrier to recovery: the consequences of stigma for the self-esteem of people with mental illnesses. *Psychiatr. Serv.* 52, 1621–1626. doi:10.1176/appi.ps.52.12.1621
- Lucas, G. M., Gratch, J., King, A., and Morency, L. P. (2014). It’s only a computer: virtual humans increase willingness to disclose. *Comput. Human Behav.* 37, 94–100. doi:10.1016/j.chb.2014.04.043
- Matsuyama, M., Bhardwaj, A., Zhao, R., Romero, O., Akoju, S., and Cassell, J. (2016). “Socially-aware animated intelligent personal assistant agent,” in *17th Annual SIGDIAL Meeting on Discourse and Dialogue*. Los Angeles.
- McLay, R. N., Deal, W. E., Murphy, J. A., Center, K. B., Kolkow, T. T., and Grieger, T. A. (2008). On-the-record screenings versus anonymous surveys in reporting PTSD. *Am. J. Psychiatry* 165, 775–776. doi:10.1176/appi.ajp.2008.07121960
- Miller, L. C., Berg, J. H., and Archer, R. L. (1983). Openers: individuals who elicit intimate self-disclosure. *J. Pers. Soc. Psychol.* 44, 1234. doi:10.1037/0022-3514.44.6.1234
- Morency, L. P., de Kok, I., and Gratch, J. (2009). A probabilistic multimodal approach for predicting listener backchannels. *Auton. Agent. Multi Agent Syst.* 20, 70–84. doi:10.1007/s10458-009-9092-y
- National Institute of Mental Health. (2010). *The Numbers Count: Mental Disorders in America*. Bethesda: National Institute of Mental Health.
- Olson, C. B., Stander, V. A., and Merrill, L. L. (2004). The influence of survey confidentiality and construct measurement in estimating rates of childhood victimization among Navy recruits. *Mil. Psychol.* 16, 53. doi:10.1207/s15327876mp1601_4
- Pickard, M. D., Roster, C. A., and Chen, Y. (2016). Revealing sensitive information in personal interviews: is self-disclosure easier with humans or avatars and under what conditions? *Comput. Human Behav.* 65, 23–30. doi:10.1016/j.chb.2016.08.004
- Qu, C., Brinkman, W.-P., Ling, Y., Wiggers, P., and Heynderickx, I. (2014). Conversations with a virtual human: synthetic emotions and human responses. *Comput. Human Behav.* 34, 58–68. doi:10.1016/j.chb.2014.01.033
- Rizzo, A. A., and Koenig, S. (in press). Is clinical virtual reality ready for primetime? *Neuropsychology*.
- Rizzo, A. A., Lange, B., Buckwalter, J. G., Forbell, E., Kim, J., Sagae, K., et al. (2011). An intelligent virtual human system for providing healthcare information and support. *Stud. Health Technol. Inform.* 163, 503–509. doi:10.1515/IJDHD.2011.046
- Rizzo, A. A., Scherer, S., DeVault, D., Gratch, J., Artstein, R., Hartholt, A., et al. (2016). Detection and computational analysis of psychological signals using a virtual human interviewing agent. *J. Pain Manag.* 9, 311–321.
- Rizzo, A. A., and Shilling, R. (in press). Clinical virtual reality tools to advance the prevention, assessment, and treatment of PTSD. *Eur. J. Psychotraumatol.*

- Schreiber, M., and McEnany, G. P. (2015). Stigma, American military personnel and mental health care: challenges from Iraq and Afghanistan. *J. Ment. Health* 24, 54–59. doi:10.3109/09638237.2014.971147
- Sebestik, J., Zelon, H., DeWitt, D., O'Reilly, J. M., and McGowan, K. (1988). "Initial experiences with CAPI," in *Proceedings of the Bureau of the Census Fourth Annual Research Conference*, Arlington, VA, 357–365.
- Serowik, K. L., Ablondi, K., Black, A. C., and Rosen, M. I. (2014). Developing a benefits counseling website for veterans using motivational interviewing techniques. *Comput. Human Behav.* 37, 26–30. doi:10.1016/j.chb.2014.03.019
- Slack, W. V., and Van Cura, L. J. (1968). Patient reaction to computer-based medical interviewing. *Comput. Biomed. Res.* 1, 527–531. doi:10.1016/0010-4809(68)90018-9
- Thomas, J. L., Wilk, J. E., Riviere, L. A., McGurk, D., Castro, C. A., and Hoge, C. W. (2010). Prevalence of mental health problems and functional impairment among active component and National Guard soldiers 3 and 12 months following combat in Iraq. *Arch. Gen. Psychiatry* 67, 614–623. doi:10.1001/archgenpsychiatry.2010.54
- Thornberry, O., Rowe, B., and Biggar, R. (1990). Use of CAPI with the U.S. National Health Interview Survey. *Paper Presented at the World Congress of Sociology, Madrid*.
- Thornicroft, G., Brohan, E., Rose, D., Sartorius, N., Leese, M., and INDIGO Study Group. (2009). Global pattern of experienced and anticipated discrimination against people with schizophrenia: a cross-sectional survey. *Lancet* 373, 408–415. doi:10.1016/S0140-6736(08)61817-6
- Tickle-Degnen, L., and Rosenthal, R. (1990). The nature of rapport and its non-verbal correlates. *Psychol. Inquiry*, 1, 285–293. doi:10.1207/s15327965pli0104_1
- van der Heijden, P. G., Van Gils, G., Bouts, J., and Hox, J. J. (2000). A comparison of randomized response, computer-assisted self-interview, and face-to-face direct questioning eliciting sensitive information in the context of welfare and unemployment benefit. *Soc. Method Res.* 28, 505–537. doi:10.1177/004912410028004005
- Vincenzen, K. (2013). Comparison of civilian trauma and combat trauma. *Vistas Online* 45, 1–8.
- Warner, C. H., Appenzeller, G. N., Grieger, T., Belenkiy, S., Breitbach, J., Parker, J., et al. (2011). Importance of anonymity to encourage honest reporting in mental health screening after combat deployment. *Arch. Gen. Psychiatry* 68, 1065–1071. doi:10.1001/archgenpsychiatry.2011.112
- Warner, C. H., Appenzeller, G. N., Mullen, K., Warner, C. M., and Grieger, T. (2008). Soldier attitudes toward mental health screening and seeking care upon return from combat. *Mil. Med.* 173, 563–569. doi:10.7205/MILMED.173.6.563
- Warner, C. H., Breitbach, J. E., Appenzeller, G. N., Yates, V., Grieger, T., and Webster, W. G. (2007). Division mental health in the new brigade combat team structure: part II. Redeployment and postdeployment. *Mil. Med.* 172, 912–917. doi:10.7205/MILMED.172.9.912
- Weisband, S., and Kiesler, S. (1996). "Self disclosure on computer forms: meta-analysis and implications," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC: ACM), 3–10.
- World Economic Forum. (2011). *The Global Economic Burden of Non-Communicable Diseases*, Cambridge: Harvard.
- World Health Organization. (2004). *The Global Burden of Disease: 2004 Update*, Geneva: World Health Organization.
- Wright, K. M., Thomas, J. L., Adler, A. B., Ness, J. W., Hoge, C. W., and Castro, C. A. (2005). Psychological screening procedures for deploying US Forces. *Mil. Med.* 170, 555–562. doi:10.7205/MILMED.170.7.555
- Zhao, R., Sinha, T., Black, A., and Cassell, J. (2016). "Socially-aware virtual agents: automatically assessing dyadic rapport from temporal patterns of behavior," in *16th International Conference on Intelligent Virtual Agents*, Los Angeles.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Lucas, Rizzo, Gratch, Scherer, Stratou, Boberg and Morency. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.