Effects of Cognitive Behavioral Stress Management Delivered by a Virtual Human, Teletherapy, and an E-Manual on Psychological and Physiological Outcomes in Adult Women: An Experimental Test

Kate Loveys 1, Michael Antoni 2, Liesje Donkin 1, Mark Sagar 3,4, William Xu 1 and Elizabeth Broadbent 1,*

1 Department of Psychological Medicine, The University of Auckland School of Medicine, 22-30 Park Avenue, Grafton, Auckland 1023, New Zealand
2 Center for Psycho-Oncology Research, The University of Miami, 5665 Ponce De Leon Drive, Coral Gables, FL 33146-0751, USA
3 Soul Machines, 1/106 Customs Street West, Auckland 1010, New Zealand
4 Auckland Bioengineering Institute, The University of Auckland, 6/70 Symonds Street, Grafton, Auckland 1010, New Zealand
* Correspondence: e.broadbent@auckland.ac.nz

Abstract: Technology may expand the reach of stress management to broader populations. However, issues with engagement can reduce intervention effectiveness. Technologies with highly social interfaces, such as virtual humans (VH), may offer advantages in this space. However, it is unclear how VH compare to telehealth and e-manuals at delivering psychological interventions. This experiment compared the effects of single laboratory session of Cognitive Behavioral Stress Management (CBSM) delivered by a VH (VH-CBSM), human telehealth (T-CBSM), and an e-manual (E-CBSM) on psychological and physiological outcomes in a community sample of stressed adult women. A pilot randomized controlled trial (RCT) with a parallel, mixed design was conducted. Adult women (M age =43.21, SD = 10.70) who self-identified as stressed were randomly allocated to VH-CBSM, T-CBSM, or E-CBSM involving one 90 min session and homework. Perceived stress, stress management skills, negative affect, optimism, relaxation, and physiological stress were measured. Mixed factorial ANOVAs and pairwise comparisons with Bonferroni correction investigated main and interaction effects of time and condition. Participants’ data (N = 38) were analysed (12 = VH-CBSM; 12 = T-CBSM; 14 = E-CBSM). Each condition significantly improved stress, negative affect, optimism, relaxation, and physiological stress over time with large effect sizes. No significant differences were found between conditions on outcomes. Overall, all three technologies showed promise for remotely delivering CBSM in a controlled setting. The findings suggest feasibility of the VH-CBSM delivery approach and support conducting a fully powered RCT to examine its effectiveness when delivering a full 10-week CBSM intervention.

Keywords: virtual human; telehealth; e-manual; stress; cognitive behavioral stress management; pilot randomized controlled trial; pilot study

1. Introduction

Virtual humans (VH) are a new form of embodied conversational agent (ECA) that are capable of social engagement [1], yet are highly scalable and accessible over an internet connection. VH include a realistic, humanlike embodiment that is typically presented on a computer, tablet, or smartphone screen, however VH can also appear in augmented or virtual reality. VH use artificial intelligence techniques such as artificial neural networks to inform their social interactions that are based on user data (e.g., emotional facial expressions, language, physiological or health data) [2]. Because of their humanlike embodiment, VH are capable of delivering verbal and non-verbal communication cues for relationship building
and emotional expression, which may help to promote engagement in the content that they deliver [3].

Technologies with highly social interfaces, such as VH, are uniquely positioned to expand the reach of psychological interventions in an engaging way. Many people face barriers to accessing psychological treatment due to issues such as a shortage of trained mental health professionals, long waitlists, high treatment costs, stigma, low health literacy, and physical restrictions (e.g., from the COVID-19 pandemic, living rurally) [4]. Increasingly, evidence is building in support of VH for remotely delivering evidence-based psychological interventions. A recent systematic review found that digital interventions with a VH were associated with significant improvements in psychological outcomes over those without a VH with a small effect size [2]. This may be because highly interactive technologies have been associated with increased intervention engagement [5]. However, the research field is still developing with relatively few trials focused on specific therapies, outcomes, or populations. Moreover, different VH may vary in functionality and acceptability.

Recently developed Digital Humans (DH) have a hyper-realistic embodiment informed by sophisticated computer-generated imagery and Hollywood film animation techniques, thus they appear very humanlike [6]. DH analyze the speech and behavioral data of users in real-time to predict their emotional state and inform their autonomous multimodal empathy responses (e.g., head nods, emotional expressions, speech, eye gaze). Due to their engaging social presence and empathetic interactions, DH seem a promising way to deliver psychological interventions but lack an evidence base to date.

**Use Case: Stress Management**

Cognitive Behavioral Stress Management (CBSM) is a psychological intervention that teaches cognitive and behavioral techniques to help people reduce their stress and improve coping skills [7]. CBSM includes techniques such as cognitive restructuring, building awareness of stress and cognitive distortions, deep breathing, and progressive muscle relaxation, amongst other techniques. CBSM has been found to improve psychological and physiological health outcomes including stress, stress management skills, distress, optimism, quality of life, serum cortisol, and immune function [8–12]. CBSM has shown benefit in a range of populations including women with breast cancer [9], men with prostate cancer [13], men with human immunodeficiency virus (HIV) [14], pregnant women [15], people with chronic fatigue syndrome [16], healthcare professionals [17], and healthy adults [18].

Therapeutic technologies with human facilitation, such as teletherapy over video- or telephone-calling, were shown to improve the accessibility of psychological therapy during the COVID-19 pandemic [19,20]. When delivered through teletherapy, CBSM has been found to significantly reduce perceived stress [21]. Similar interventions have also been shown to be effective in people with cancer when delivered through a collaborative care program involving a website and remote contact with a therapist [22]. Although promising, teletherapy services are affected by the shortage of trained mental health professionals, and they may be more costly and less scalable compared to other digital health technologies without human facilitation.

Technology interventions without facilitation, such as self-guided websites and smartphone applications, are another promising option for intervention delivery. Research evaluating CBSM delivery in a self-guided format is limited, however one study found that a smartphone application was a feasible and acceptable way to deliver CBSM to people living with HIV [23]. Moreover, preliminary evidence of effectiveness was found. Other research has found that CBSM delivered by an online workbook was effective for improving coping self-efficacy, mood, and post-traumatic stress symptoms in people with breast cancer [24], and that the delivery format can maintain treatment fidelity [25]. A larger number of trials have looked at the delivery of other cognitive behavioral or mindfulness interventions using self-guided technologies, and have shown that they can be effective and acceptable formats to support stress management [26,27]. However, there may be
low engagement with self-guided technology interventions after first use and outside of clinical trial contexts [28,29]. This may be because they lack the social accountability and engagement that a human therapist would typically provide [5].

VH may increase intervention engagement compared to self-guided technologies because they are more humanlike and socially interactive. However, to date, no studies have investigated the effects of CBSM when delivered by a VH. There is limited research investigating the effectiveness of other stress management interventions when delivered by a VH, however several studies have found promising results. When used to deliver a holistic stress management program, a VH was found to significantly improve health behaviours compared to a patient education sheet with the same content [30]. Another study evaluated a VH that provided home-based coaching between group stress management sessions for people with chronic pain and depression [31]. In this study, supplementary support from a VH was found to significantly improve stress management behaviours compared to usual care. More recent work found that VH may be acceptable and feasible for use in delivering psychoeducation and deep breathing exercises for stress to at-risk younger and older adults during the COVID-19 pandemic [32]. However, there is limited evidence on how VH compare to other common eHealth technologies such as telehealth and self-guided e-manuals at delivering interventions and additional studies are needed. Moreover, it is unclear whether the different delivery modalities affect practice of psychological exercises beyond the treatment session.

We conducted an experiment to compare the effects of a VH (VH-CBSM), telehealth (T-CBSM), and a self-guided e-manual (E-CBSM) delivering one session of CBSM on psychological and physiological stress markers, in a community sample of adult women in distress. A sample of women were selected to remove gender as a confound, as women have been shown to experience greater stress than men on average [33], including during the COVID-19 pandemic when this research was conducted [34]. Results regarding the feasibility and acceptability of the technologies from this study are presented elsewhere [35]. Because all three conditions involved one session of CBSM, we hypothesized that women assigned to any of the three delivery methods would show some improvements in psychological and physiological stress markers. In addition, it was hypothesized that those receiving either DH-CBSM or T-CBSM would show greater improvements than those assigned to E-CBSM due to the greater social facilitation offered.

2. Materials and Methods

A pilot study with a parallel, mixed design was conducted. A community sample of 43 adult women who were fluent in English and self-identified as stressed were randomly allocated to receive one 90 min session of CBSM in a controlled laboratory setting using one of three delivery methods: a virtual human (VH-CBSM), a human teletherapist over video call (T-CBSM), or a self-guided e-manual (E-CBSM) (see Figure 1). The session took place in a private clinic room at the University of Auckland Clinical Research Centre, which enabled collection of physiological data. Each technology delivered identical content that covered psychoeducation about stress awareness and a deep breathing exercise adapted from a CBSM treatment manual [36]. The VH was a Digital Human by Soul Machines (Auckland, New Zealand). The VH engaged in rapport building behaviours (e.g., emotional expressiveness, empathy) that were informed by its multimodal emotion recognition and response system. The system classified participant speech and facial behaviours for emotions in real-time from audiovisual data collected during the interaction. The teletherapy session was hosted in Zoom videocconferencing software (Zoom Video Communications Inc., San Jose, CA, USA) by a Masters level trainee health psychologist who was supervised by a senior clinical and health psychologist (LD). The e-manual was hosted in Qualtrics (Qualtrics, Provo, UT, USA).
After the session, participants were asked to complete daily homework over two weeks which involved a deep breathing exercise and watching three educational videos about stress on a website. Participants received a $30 shopping voucher for their time.

The protocol was prospectively registered with the Australia New Zealand Clinical Trials Registry (ANZCTR) on 28/08/2020 (registration no. ACTRN12620000859987). The University of Auckland Human Participants Ethics Committee provided approval on 17/12/2019 (reference no. 024085).

2.1. Measures

2.1.1. Perceived Stress

The 10-item Perceived Stress Scale (PSS) evaluated the extent to which events over the past month (at baseline) or past two weeks (at the two-week follow-up) were appraised as stressful [37]. Responses could range from 0 (‘never’) to 4 (‘very often’). Total scores range from 0 to 40, with higher scores indicating greater perceived stress. The PSS has demonstrated good test–retest reliability (\(\alpha > 0.70\)), convergent and divergent validity [38]. The PSS showed good internal consistency reliability in this sample (\(\alpha = 0.79\)).

Present moment stress and relaxation at baseline, post-intervention, and at the two-week follow-up assessment were measured using two 100 mm Visual Analogue Scales (VAS) from 0 (‘not at all’) to 100 (‘extremely’) [39]. Participants also completed a daily stress assessment during the two weeks after their appointment using a 100 mm VAS from 0 (‘not at all’) to 100 (‘extremely’), which was delivered by email each morning. However, due to a low response rate to the daily stress measure, these data were not analyzed (47%, 18 of 38 participants provided complete data).

2.1.2. Negative Affect

Negative affect was measured at baseline and at the two-week follow-up assessment using a 10-item composite measure comprised of the anxiety, depression, and anger subscales from the Profile of Mood States (POMS) scale [40]. This composite measure has been used in prior CBSM research with adult women [41]. Participants indicated the extent to which they felt the emotion listed over the past month (at baseline) or the past two weeks (at the two-week follow-up) using a 5-point scale (0 = ‘not at all’–4 = ‘extremely’). The POMS was shown to have good internal consistency reliability in our sample (\(\alpha = 0.88\)).

2.1.3. Optimism

Optimism was assessed using the Revised Life Orientation Test (LOT-R) at baseline and at the two-week follow-up assessment [42]. Total scores could range from 0 to 24, and higher scores indicated greater optimism. This scale has been used to measure changes in optimism in prior CBSM studies with good internal reliability, test–retest reliability, and divergent validity with distress and depression [41]. Adequate internal consistency reliability was found in our sample (\(\alpha = 0.75\)).
2.1.4. Perceived Stress Management Skills

Participants’ perceived ability to perform the stress management techniques covered in the CBSM session was measured using two subscales from the Measure of Current Status (MOCS) scale: relaxation and awareness of tension [9]. Together, the two subscales totaled five items. Responses could range from 0 (‘I cannot do this at all’) to 4 (‘I can do this extremely well’). Perceived stress management skills were reported post-intervention and at the two-week follow-up (not baseline). The scale demonstrated adequate internal consistency reliability in our sample (α = 0.76).

2.1.5. Physiological Stress

Electrodermal activity (EDA) and skin temperature were measured during the session using a wrist-worn sensor device (Empatica E4; Empatica, Boston, MA, USA). Data were collected continuously and began after a 10 min rest period at the start of the session, and finished at the end of the session. Artifacts were removed that may confound that data (e.g., movement) using prior methods [43]. EDA was automatically processed using a validated algorithm that calculated counts of skin conductance responses (or “peaks”) in the raw data [44]. A lower number of skin conductance peaks indicated lower physiological stress [45]. Skin temperature was measured in degrees Celsius. Higher skin temperature indicated reduced sympathetic arousal (i.e., lower physiological stress) [46]. Average number of skin conductance responses and skin temperature scores were calculated at: baseline (1 min of data from the start of the recording), mid-point of the session (1 min of data at the end of the psychoeducation session and before the beginning of the deep breathing exercise), and end of the session (1 min of data before the end of the deep breathing exercise). The mid-point of the session timepoint enabled evaluation of the physiological data after the first part of the session (psychoeducation and cognitive exercises), and prior to the start of the second part (deep breathing exercise).

2.2. Data Analysis

Data were analyzed using IBM SPSS software (version 27, Armonk, NY, USA). Mean imputation handled missing data where applicable. According to Little’s MCAR test, the data were missing completely at random (p = 1.000). Baseline sample characteristics were compared between groups with one-way ANOVA and chi-square tests. A series of mixed factorial ANOVAs were conducted to investigate the main and interaction effects of time and condition on psychological and physiological outcomes. Data were checked for violations of test assumptions. Greenhouse-Geisser values were reported for within-subjects effects where sphericity assumptions were violated (general stress, distress, optimism, stress management skills, electrodermal activity, skin temperature). Otherwise, test assumptions were met. Pairwise comparisons with Bonferroni correction were conducted as follow-up tests.

3. Results
3.1. Participants

43 participants were recruited, however 38 participants’ data were analysed. Reasons for dropout and exclusion, and participant flow through the study are presented in a CONSORT diagram in Appendix A. Participants with analyzable data were female adults (mean age = 43.21, SD = 10.70), who were highly educated, with 60.5% (23 of 38) holding a post-graduate degree. Most participants were of New Zealand European (63.2%; 24 of 38) or Asian ethnicity (23.7%; 9 of 38). Just under one third (28.9%; 11 of 38) of participants reported a mental health diagnosis (4 = major depression, 5 = generalized anxiety; 2 = depression and anxiety). On average, participants reported a moderate degree of stress at baseline based on PSS scores (M = 21.92, SD = 5.23). There were no significant group differences in demographics (all ps > 0.119) or psychological variables at baseline (all ps > 0.256). Baseline demographics and psychological variables are presented in Appendix B.
3.2. State Stress (100 mm VAS)

There was a large, significant main effect of time on state levels of stress \((F(2, 62) = 33.02, p < 0.001, \eta^2_p = 0.52)\). Pairwise comparisons revealed that compared to baseline, participants reported significantly lower stress immediately after the CBSM session (Baseline: \(M = 61.84, SE = 3.53, 95\% \text{ CI } [54.64, 69.04], \text{range} = 10–100\) vs. Post-intervention: \(M = 28.07, SE = 3.38, 95\% \text{ CI } [21.18, 34.97], \text{range} = 0–75, F(2, 30) = 43.14, p < 0.001, \eta^2_p = 0.74\), and at two-weeks follow-up across all conditions (\(M = 45.43, SE = 3.68, 95\% \text{ CI } [37.91, 52.94], \text{range} = 5–85, F(2, 30) = 43.14, p = 0.005, \eta^2_p = 0.74\)) (see Figure 2). Participants reported significantly less stress at the end of the CBSM session compared to two-weeks later (\(F(2, 62) = 43.14, p < 0.001, \eta^2_p = 0.74\)).

There was no significant main effect of condition \((F(2, 31) = 2.13, p = 0.136, \eta^2_p = 0.12)\), nor was there a significant interaction effect of time and condition on state stress \((F(4, 62) = 0.29, p = 0.881, \eta^2_p = 0.02)\).

3.3. State Relaxation (100 mm VAS)

There was a large, significant main effect of time on relaxation \((F(2, 60) = 11.47, p < 0.001, \eta^2_p = 0.28)\). Pairwise comparisons showed that overall, participants were significantly more relaxed at the end of the CBSM session (M = 42.55, SE = 3.54, 95% CI [35.33, 49.77], range = 5–100, compared to baseline (M = 67.17, SE = 4.12, 95% CI [58.75, 75.58], range = 10–84, F(2, 29) = 10.60, p < 0.001, \eta^2_p = 0.42), and compared to the two week follow-up assessment (M = 49.64, SE = 4.40, 95% CI [40.65, 58.63], range = 3–95, F(2, 29) = 10.60, p = 0.006, \eta^2_p = 0.42) (see Figure 3).
There was no significant main effect of condition (F(2, 30) = 0.73, \( \eta^2_p = 0.05 \)), nor interaction effect of time and condition on state relaxation (F(4, 60) = 1.32, \( p = 0.273, \eta^2_p = 0.08 \)).

### 3.4. Perceived Stress Scale

There was a large, significant main effect of time on PSS scores (F(1, 30) = 30.43, \( p < 0.001, \eta^2_p = 0.50 \)). Pairwise comparisons revealed that overall, participants reported significantly lower perceived stress at the two week follow-up compared to baseline (M = 16.80, SE = 1.20, 95% CI [14.34, 19.25], range = 6–36 vs. M = 22.15, SE = 0.98, 95% CI [20.16, 24.14], range = 10–33). There was no significant main effect of condition (F(2, 30) = 0.56, \( p = 0.578, \eta^2_p = 0.04 \)), nor was there a significant interaction effect between condition and time on the PSS (F(2, 30) = 0.81, \( p = 0.455, \eta^2_p = 0.05 \)).

### 3.5. Negative Affect

There was a large, significant main effect of time on negative affect (NA) (F(1, 30) = 26.28, \( p < 0.001, \eta^2_p = 0.47 \)). Participants reported significantly lower NA at the two-week follow-up compared to baseline across all conditions (M = 3.89, SE = 0.44, 95% CI [3.01, 4.78], range = 1–11.67 vs. M = 5.79, SE = 0.43, 95% CI [4.92, 6.65], range = 1–11). There was no significant main effect of condition (F(2, 30) = 0.49, \( p = 0.619, \eta^2_p = 0.03 \)), or interaction effect of condition and time on NA (F(2, 30) = 1.68, \( p = 0.203, \eta^2_p = 0.10 \)).

### 3.6. Optimism

There was a large, significant main effect of time on optimism (F(1, 30) = 10.78, \( p = 0.003, \eta^2_p = 0.26 \)). Pairwise comparisons showed that participants reported significantly greater optimism at the two week follow-up assessment (M = 16.04, SE = 0.77, 95% CI [14.48, 17.60], range = 6–24) compared to baseline (M = 14.60, SE = 0.80, 95% CI [12.97, 16.23], range = 3–22). There was no significant main effect of condition (F(2, 30) = 2.06, \( p = 0.145, \eta^2_p = 0.08 \)).
\( \eta^2_p = 0.12 \), or significant interaction effect of condition and time on optimism (\( F_{(2, 30)} = 0.22, p = 0.807, \eta^2_p = 0.01 \)).

### 3.7. Perceived Stress Management Skills

There was a trend towards a significant main effect of time on awareness of tension skills, with a moderate effect size (\( F_{(1, 31)} = 3.06, p = 0.090, \eta^2_p = 0.09 \)). Participants reported greater awareness of tension skills at the two week follow-up assessment (\( M = 7.21, \text{SE} = 0.45, 95\% \text{CI} [6.30, 8.13], \text{range} = 1–12 \)), compared to immediately after the therapy session (\( M = 6.49, \text{SE} = 0.43, 95\% \text{CI} [5.61, 7.36], \text{range} = 1–11 \)). There was no significant main effect of condition (\( F_{(2, 31)} = 0.22, p = 0.802, \eta^2_p = 0.01 \)), or interaction effect between time and condition on perceived awareness of tension skills (\( F_{(2, 31)} = 0.20, p = 0.816, \eta^2_p = 0.01 \)).

Participants reported moderate relaxation skills at the end of the CBSM session (\( M = 3.82, \text{SE} = 0.36, 95\% \text{CI} [3.07, 4.56], \text{range} = 0–8 \)). There was no significant main effect of time (\( F_{(2, 31)} = 2.52, p = 0.123, \eta^2_p = 0.08 \)). However, there was a non-significant improvement in the overall sample means for relaxation skills between the end of the session and two-week follow-up assessments, with a moderate effect size (\( M = 3.82, \text{SE} = 0.36, 95\% \text{CI} [3.07, 4.56], \text{range} = 0–8 \) vs. \( M = 4.41, \text{SE} = 0.35, 95\% \text{CI} [3.70, 5.11], \text{range} = 2–8 \), \( F_{(2, 31)} = 2.52, p = 0.123, \eta^2_p = 0.08 \)). There was no significant main effect of condition (\( F_{(2, 31)} = 0.24, p = 0.787, \eta^2_p = 0.02 \)), or interaction effect of condition and time on perceived relaxation skills (\( F_{(2, 31)} = 1.08, p = 0.533, \eta^2_p = 0.07 \)).

### 3.8. Electrodermal Activity

Analyses revealed a large, significant main effect of time on electrodermal activity (EDA) (\( F_{(1.39, 42.94)} = 8.49, p = 0.003, \eta^2_p = 0.22 \)). Pairwise comparisons showed that compared to baseline, EDA was significantly reduced at the mid-point of the session (\( M = 1.07, \text{SE} = 0.23, 95\% \text{CI} [0.60, 1.55], \text{range} = 0–5 \) vs. \( M = 0.34, \text{SE} = 0.16, 95\% \text{CI} [0.03, 0.66], \text{range} = 0–4 \)), and at the end of the session across conditions (\( M = 0.26, \text{SE} = 0.12, 95\% \text{CI} [0.02, 0.50], \text{range} = 0–3 \), \( F_{(2, 30)} = 5.26, p = 0.007, \eta^2_p = 0.26 \)). There was no significant difference in EDA from the mid-point to the end of the session (\( p = 1.00 \)) (see Figure 4).

There was a trend towards a significant main effect of condition on EDA (\( F_{(2, 31)} = 3.13, p = 0.058, \eta^2_p = 0.17 \)). Pairwise comparisons revealed a trend towards a significant difference between the E-CBSM and VH-CBSM conditions (\( F_{(2, 31)} = 3.13, p = 0.057, \eta^2_p = 0.17 \) (see Figure 4). Participants in the E-CBSM condition had greater EDA over the CBSM session compared to participants in the VH-CBSM condition (\( M = 0.89, \text{SE} = 0.21, 95\% \text{CI} [0.47, 1.31], \text{range} = 0–4 \) vs. \( M = 0.15, \text{SE} = 0.22, 95\% \text{CI} [−0.29, 0.59], \text{range} = 0–2 \)). No other trends or significant effects were found between conditions in EDA (all \( ps > 0.364 \)). There was no significant interaction effect between time and condition on EDA (\( F_{(2, 72, 42.94)} = 2.21, p = 0.106, \eta^2_p = 0.13 \)).

### 3.9. Skin Temperature

There was a large, significant main effect of time on skin temperature (\( F_{(1.22, 37.86)} = 41.67, p = <0.001, \eta^2_p = 0.57 \)). Pairwise comparisons revealed that compared to baseline, participants had significantly greater skin temperature (i.e., reduced sympathetic arousal) at the mid-point of the session after completing the cognitive exercises (\( M = 31.34, \text{SE} = 0.26, 95\% \text{CI} [30.82, 31.87], \text{range} = 27.15–33.93 \) vs. \( M = 32.77, \text{SE} = 0.29, 95\% \text{CI} [32.19, 33.36], \text{range} = 28.52–35.73 \), \( F_{(2, 30)} = 42.54, p = <0.001, \eta^2_p = 0.74 \)) and at the end of the session after the deep breathing exercise across all conditions (\( M = 32.71, \text{SE} = 0.33, 95\% \text{CI} [32.05, 33.37], \text{range} = 28.32–35.68 \), \( F_{(2, 30)} = 42.54, p = <0.001, \eta^2_p = 0.74 \)). There was no significant difference in skin temperature scores from the mid-point to the end of the session (\( p = 1.00 \)) (see Figure 5).
There was no significant main effect of condition ($F_{(2, 31)} = 1.028, p = 0.370, \eta^2_p = 0.06$), or significant interaction effect of condition and time on skin temperature ($F_{(2, 44, 37.86)} = 0.31, p = 0.778, \eta^2_p = 0.02$).
4. Discussion

Technology may help to expand access to CBSM by overcoming some of the barriers to care associated with in-person treatment, such as shortages of trained mental health professionals, physical restrictions, stigma, and cost [47,48]. However, to date, it is an open question as to which technologies are best suited to delivering psychological interventions. This study compared different experimental manipulations of CBSM delivery by a VH, teletherapy, and a self-guided e-manual in a community sample of distressed adult women. We found that all three venues were associated with significant improvements in psychological and physiological markers with large effect sizes, many of which appeared to be retained over a 2-week follow-up period. This preliminary finding is in keeping with the broader CBSM literature showing that CBSM is an effective intervention for improving psychological and physiological outcomes [49], though this is the first time that such effects were demonstrated using a single session of CBSM. The CBSM session delivered deep breathing exercises which have been shown to reduce psychological and physiological stress [50]. However, this study is the first to compare the effects of a VH, telehealth, and an e-manual at delivering CBSM.

There were no clear differences in the effects obtained across experimental conditions. It is possible that differences in effects between groups could emerge in a fully powered RCT using a clinical sample and/or a longer intervention. Prior research suggests that both self-guided and human-facilitated interventions are effective for improving stress, however the effect sizes for psychological interventions with human facilitation tend to be larger than those delivered via self-guided technologies (e.g., websites, smartphone applications) [51]. Albeit, some meta-analyses did not find any significant difference in the effect sizes of self-guided and human-facilitated interventions, which may be due to intervention and population factors [52,53]. VH fall somewhere in between these two categories as they are self-guided technologies with a highly social interface, creating a similar experience to a human-facilitated telepsychology intervention. A recent meta-analysis found that VH are typically more effective than self-guided e-health interventions across a range of health applications, with a small effect size [2], however, their comparison to telehealth interventions is understudied. Moreover, the meta-analysis was based on a small number of studies that focused on stress management and other applications. Additional research is needed that compares these technologies at delivering stress management interventions.

Differences in technology engagement may also have implications for intervention effectiveness. Data from the two-week follow-up period gives an indication of the effects of the session plus engagement in the homework exercises. There were no significant differences between the conditions on homework engagement in this sample as reported elsewhere [35]. In a longer trial with a full 10-week CBSM intervention accessed from home, we may have seen variation in engagement between the three delivery approaches over time which may impact their effectiveness. According to prior research, we might expect engagement with the self-guided e-manual to drop off earlier in the trial compared to the telehealth and VH conditions that use social engagement strategies [5]. As the research field is developing, it is unclear how engagement with the telehealth and VH conditions would compare over a full-length intervention.

Limitations and Future Research

There were several limitations of this study which indicate areas for future research. As this was a pilot study, it was not powered to detect significant group differences, however the within-group changes can be used to inform sample size calculations for a fully powered RCT. Another limitation is that the sample population consisted of predominantly White or Asian, highly educated women. It is unclear whether the results would be replicated in a more diverse sample. However, other research has found promising results for VH [2], telehealth [54], and self-guided internet interventions in broader populations [51]. As this study had an all-female sample, the findings only generalize to females. However, CBSM has been shown to be beneficial for males in other research [13,14]. Although this
study contained an ethnically diverse sample, future studies should evaluate whether the results generalize to other cultures, countries, contexts, and genders. Moreover, our participants were moderately stressed at baseline and it is unknown whether the findings would generalize to severely distressed groups or those with serious medical conditions undergoing treatments. CBSM has been commonly used with women with breast cancer and HIV-positive people who may be more distressed than a community sample [49]. Further research is needed to examine how well these technologies compare at delivering CBSM to clinical samples experiencing greater distress. Another consideration is that this trial tested delivery of one CBSM session and research is needed to compare the performance of these technologies at delivering a full 10-week intervention program. Thus, additional CBSM modules should be adapted for VH, telehealth, and e-manual delivery. A future trial could assess stress management skills at baseline to evaluate within-person changes and check for baseline group differences. This study did not control for time of day or day of week effects as participants took part across all days of the week and at different times of day, thus any effects of these variables should have been counteracted. However, future research should be mindful of these possible confounds.

5. Conclusions

VH, telehealth, and self-guided e-manuals have the potential to increase the accessibility of stress management interventions such as CBSM to broader populations, however additional trials are needed to establish effectiveness. All three technologies were found to improve psychological and physiological outcomes with large effect sizes in a community sample of adult women experiencing stress. The results support conducting a fully powered RCT to compare the effectiveness of these delivery technologies over a full 10-week CBSM intervention in a clinical population.


Funding: This research received no external funding, however Soul Machines (Auckland, New Zealand) provided the virtual human used in the research.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Human Participants Ethics Committee of the University of Auckland on 17/12/2019 (reference no. 024085).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ethical reasons; participants did not provide consent for the data to be shared publicly.

Acknowledgments: The authors would like to thank Soul Machines (Auckland, New Zealand) for providing the virtual human used in the research, and Ayush Nigam, Roisin Kelly (www.rkcreator.com), Nikita Kirkaldy, and Sarah Rankin for their involvement in the research.

Conflicts of Interest: MS is the Co-Founder and Chief Scientific Officer of Soul Machines, which supported K.L. with a PhD stipend during the research and currently employs K.L., and contracts E.B. for consultancy work. M.A. is the inventor of CBSM (UMIP-483), and is a paid consultant for Blue Note Therapeutics and Atlantis Healthcare, two software companies that develop digital healthcare solutions. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Trial Registration: Australia New Zealand Clinical Trials Registry (ANZCTR): ACTRN12620000859987.
Appendix A

Conflicts of Interest: MS is the Co-Founder and Chief Scientific Officer of Soul Machines, which supported K.L. with a PhD stipend during the research and currently employs K.L., and contracts E.B. for consultancy work. M.A. is the inventor of CBSM (UMIP-483), and is a paid consultant for Blue Note Therapeutics and Atlantis Healthcare, two software companies that develop digital healthcare solutions. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Trial Registration: Australia New Zealand Clinical Trials Registry (ANZCTR): ACTRN12620000859987

Figure A1. CONSORT Flow Diagram.
Appendix B

Table A1. Baseline demographic and psychological variables.

<table>
<thead>
<tr>
<th></th>
<th>Total (N = 38)</th>
<th>E-CBSM (n = 14)</th>
<th>VH-CBSM (n = 12)</th>
<th>T-CBSM (n = 12)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (M, SD)</strong></td>
<td>43.21 (10.70)</td>
<td>44.71 (11.93)</td>
<td>39.67 (7.97)</td>
<td>45.00 (11.55)</td>
</tr>
<tr>
<td><strong>Ethnicity:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian (n, %)</td>
<td>24 (63.2%)</td>
<td>8 (57.1%)</td>
<td>8 (66.7%)</td>
<td>8 (66.7%)</td>
</tr>
<tr>
<td>Asian (n, %)</td>
<td>9 (23.7%)</td>
<td>5 (35.7%)</td>
<td>2 (16.7%)</td>
<td>2 (16.7%)</td>
</tr>
<tr>
<td>Māori (n, %)</td>
<td>3 (7.9%)</td>
<td>1 (7.1%)</td>
<td>2 (16.7%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Middle Eastern/Latin American/African (n, %)</td>
<td>1 (2.6%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>1 (8.3%)</td>
</tr>
<tr>
<td>Other (n, %)</td>
<td>1 (2.6%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>1 (8.3%)</td>
</tr>
<tr>
<td><strong>Education level:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school or less (n, %)</td>
<td>1 (2.6%)</td>
<td>1 (7.1%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Trade qualification (n, %)</td>
<td>2 (5.3%)</td>
<td>1 (7.1%)</td>
<td>1 (8.3%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Undergraduate degree (n, %)</td>
<td>12 (31.6%)</td>
<td>6 (42.9%)</td>
<td>3 (25.0%)</td>
<td>3 (25.0%)</td>
</tr>
<tr>
<td>Postgraduate degree (n, %)</td>
<td>23 (60.5%)</td>
<td>6 (42.9%)</td>
<td>8 (66.7%)</td>
<td>9 (75.0%)</td>
</tr>
<tr>
<td><strong>Marital status:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single (n, %)</td>
<td>8 (21.1%)</td>
<td>6 (42.9%)</td>
<td>1 (8.3%)</td>
<td>1 (8.3%)</td>
</tr>
<tr>
<td>Relationship (n, %)</td>
<td>5 (13.2%)</td>
<td>1 (7.1%)</td>
<td>3 (25.0%)</td>
<td>1 (8.3%)</td>
</tr>
<tr>
<td>Married/ living with partner (n, %)</td>
<td>21 (55.3%)</td>
<td>5 (35.7%)</td>
<td>7 (58.3%)</td>
<td>9 (75.0%)</td>
</tr>
<tr>
<td>Separated/ divorced (n, %)</td>
<td>4 (10.5%)</td>
<td>2 (14.3%)</td>
<td>1 (8.3%)</td>
<td>1 (8.3%)</td>
</tr>
<tr>
<td><strong>Work status:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time (n, %)</td>
<td>27 (71.1%)</td>
<td>10 (71.4%)</td>
<td>9 (75.0%)</td>
<td>8 (66.7%)</td>
</tr>
<tr>
<td>Part-time (n, %)</td>
<td>6 (15.8%)</td>
<td>3 (21.4%)</td>
<td>2 (16.7%)</td>
<td>1 (8.3%)</td>
</tr>
<tr>
<td>Beneficiary (n, %)</td>
<td>1 (2.6%)</td>
<td>1 (7.1%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Unemployed (n, %)</td>
<td>4 (10.5%)</td>
<td>0 (0.0%)</td>
<td>1 (8.3%)</td>
<td>3 (25.0%)</td>
</tr>
<tr>
<td><strong>Mental health:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least one mental health diagnosis (n, %)</td>
<td>11 (28.9%)</td>
<td>4 (28.6%)</td>
<td>3 (25.0%)</td>
<td>4 (33.3%)</td>
</tr>
</tbody>
</table>

References


35. Loveys, K.; Antoni, M.; Donkin, L.; Sagar, M.; Broadbent, E. Comparing the feasibility and acceptability of teletherapy, a virtual human, and an e-manual at delivering a stress manage-ment intervention to distressed adult women: A pilot randomized controlled trial. Department of Psychological Medicine, The University of Auckland School of Medicine, 22-30 Park Avenue, Grafton, Auckland 1023, New Zealand. 2022; manuscript in submission.


