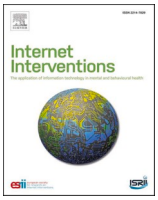
Internet Interventions 30 (2022) 100580

Contents lists available at ScienceDirect

Internet Interventions

journal homepage: www.elsevier.com/locate/invent



Momentary effects of Temstem, an app for voice-hearing individuals: Results from naturalistic data from 1048 users

Alyssa Jongeneel [a,b,*], Ilan Libedinsky [a], Anouk Reinbergen [b], Nynke Tromp [c], Philippe Delespaul [d,e], Heleen Riper [a,f], Mark van der Gaag [a], David van den Berg [a,b]

[a] Department of Clinical, Neuro- and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Van der Boechorststraat 7, 1081 BT Amsterdam, Netherlands
[b] Mark van der Gaag Research Centre, Parnassia Psychiatric Institute, Zoutkeetsingel 40, 2512 HN Den Haag, Netherlands
[c] Department of Industrial Design, Delft University of Technology, Landbergstraat 15, 2628 CE Delft, Netherlands
[d] Department of Psychiatry and Neuropsychology, University of Maastricht, Maastricht, Netherlands
[e] Mondriaan Mental Health Trust, Maastricht and Heerlen, Netherlands
[f] Department of Research and Innovation, GGZ InGeest Specialized Mental Health Care, Amsterdam, Netherlands

ARTICLE INFO

Keywords:
eHealth
App
Voice hearing
Auditory verbal hallucinations
Naturalistic data
User data

ABSTRACT

Background: Temstem is a mobile application developed in cooperation with voice-hearing persons to help them cope with distressing voices. After psychoeducation about voice hearing, Temstem offers two functions: Silencing is a mode designed to inhibit voice activity through the processing of incompatible language; the Challenging mode introduces dual tasking (as used in eye movement desensitisation and reprocessing) designed to reduce the emotionality and vividness of a voice memory. Two different language games, Lingo Tapper and Word Link, are provided, containing both functions. This study aimed to explore the momentary effects of Temstem on voice-hearing distress, emotionality and vividness in a naturalistic sample of voice-hearing app users.

Method: Temstem is freely available in the Netherlands. We collected data through the app from 1048 individual users who had given informed consent for the study. We assessed changes in pre- and post-session scores on distress, emotionality and vividness, and we evaluated differences in outcomes between the games and whether effects remained stable over multiple sessions.

Results: Users had been hearing voices for an average of 4.95 years; 79 % had been informed about Temstem by a mental health therapist or coach. After a Silencing session, voice-hearing distress was reduced, $t(958) = 27.12$, $p < .001$, $d = 0.49$; the degree of reduction remained stable after repeated use, $F(1, 7905.57) = 1.91$, $p = .167$. After a Challenging session, emotionality, $t(651) = 23.16$, $p < .001$, $d = 0.74$, and the vividness of voice memories were reduced, $t(651) = 22.20$, $p < .001$, $d = 0.71$; both diminished slightly with frequent use, $F(1, 2222.86) = 7.21$, $p < .05$; $F(1, 2289.92) = 4.25$, $p < .05$. In comparison with Lingo Tapper, larger reductions were seen for a Word Link game: emotionality $t(226) = 2.88$, $p < .005$, $d = 0.21$; vividness $t(226) = 2.29$, $p < .05$, $d = 0.17$.

Discussion: In this heterogeneous sample of voice-hearing individuals, Temstem appeared to be a promising coping tool; momentary voice-hearing distress and the emotionality and vividness of voice statements were reduced after a Temstem session. Despite important limitations and the need for more research, naturalistic studies of user app data may yield interesting and generalisable findings.

1. Introduction

Auditory hallucinations (AH) are a common experience, with lifetime prevalences of 9.6 % in the general population and 12.5 % in children and adolescents (Maijer et al., 2018). Auditory verbal hallucinations (AVHs) are prevalent in psychotic disorders, affecting 75 % of people with lifetime schizophrenia (Bauer et al., 2011). AVHs can be highly distressing, with negative affect and anxiety associated with elevated levels of distress (Johns et al., 2014).

Treatment of AVHs generally includes antipsychotic medication, and

* Corresponding author at: Department of Clinical, Neuro- and Developmental Psychology, Van der Boechorststraat 7, 1081 BT Amsterdam, Netherlands.
E-mail address: m.j.jongeneel@vu.nl (A. Jongeneel).

https://doi.org/10.1016/j.invent.2022.100580
Received 2 March 2022; Received in revised form 13 September 2022; Accepted 4 October 2022
Available online 5 October 2022

medication effects are often evaluated in terms of a cluster of positive symptoms, not AVHs specifically. Medications have variable effectiveness in reducing positive symptoms, with effect sizes ranging from small to large, but they are often not taken as prescribed due to the side-effects and long-term risks involved (Huhn et al., 2019; Sommer et al., 2012). Cognitive-behavioural therapy for psychosis (CBTp) varies in effectiveness, showing stable small to medium effect sizes, but it is hard to implement on a large scale due to training costs and time investment for therapists (Turner et al., 2020). A need therefore exists for additional, low-threshold interventions for individuals distressed by AVHs.

Mental health care is increasingly employing electronic health (eHealth) interventions (Ben-Zeev, 2012; Cuijpers et al., 2017). eHealth methods have been shown to be feasible and acceptable for use in treating people with AVHs (Thomas et al., 2019). Smartphone apps in particular have the advantage of being available at any place and time, enabling direct intervention when needed, and thus enhancing the self-management of the user (Anderson and Funnell, 2005; Calabretta, 2002). For these reasons, the Temstem app was developed for and with people with distressing voices (Jongeneel et al., 2018). Temstem intervenes to reduce AVH distress through two functions – Silencing and Challenging – each using two language game options (for descriptions, see Section 2.2, 'Procedure'). The functions are based on the mechanisms of therapeutic action described below.

First, users may experience more control over AVHs by using the Silencing (*inhibition*) function of Temstem: while hearing voices, users play a language game to temporarily suppress the voices. This function was derived from the finding that many people cannot hear voices while actively producing incompatible sounds or language; humming and gargling, for example, interrupted ongoing hallucinations, while biting the tongue, opening the mouth, clenching a fist or raising the eyebrows had no such effect (Erickson and Gustafson, 1968; Green and Kinsbourne, 1989). Blocking the vocal chords by widely opening the mouth did interrupt voices, while firmly closing one's eyes or clenching a fist did not (Bick and Kinsbourne, 1987). Especially meaningful stimuli such as real words are effective in inhibiting auditory hallucinations (Farhall et al., 2007; Gallagher et al., 1994a). Research has also shown that activating the language production and motor areas in the brain suppresses AVHs (Gallagher et al., 1994b, 1995). In short, the principle of Silencing is based on the inhibition of voice activity by introducing an incompatible language production or motor language task.

Second, the Challenging (*dual tasking*) function intervenes to reduce the salience and frequency of AVHs. A recent study by our group, involving a sample of voice-hearing individuals in specialised treatment, showed that Challenging – that is, taxing the working memory with a language game during recall of a negative voice statement – significantly reduced the emotionality, vividness and credibility of the voice statement (Jongeneel et al., 2020). Similarly, Matthijssen et al. (2019) found that dual tasking reduced emotionality in auditory hallucination memories in a voice-hearing sample. Such an effect may potentially generalise to a decrease in the frequency and distress of AVHs in daily life – attributable, for example, to an increased sense of control over the AVHs, which has been found to be an important factor in reducing voice-hearing distress (Daalman et al., 2011). In short, the principle of Challenging is to reduce the emotionality and vividness of a voice-hearing memory through dual tasking.

Temstem was and still is freely available in the Netherlands; during a certain period of time, we recorded the app data of the users, providing that they gave informed consent for the use of their data for research purposes. We analysed this self-reported app data in order to assess the following questions:

Silencing (inhibition)

1a) Is there a momentary effect of Silencing sessions on voice-hearing distress?
1b) Is there a difference between the Lingo Tapper and Word Link games in terms of momentary effect on voice-hearing distress?
1c) Are effects on voice-hearing distress sustained in multiple use?

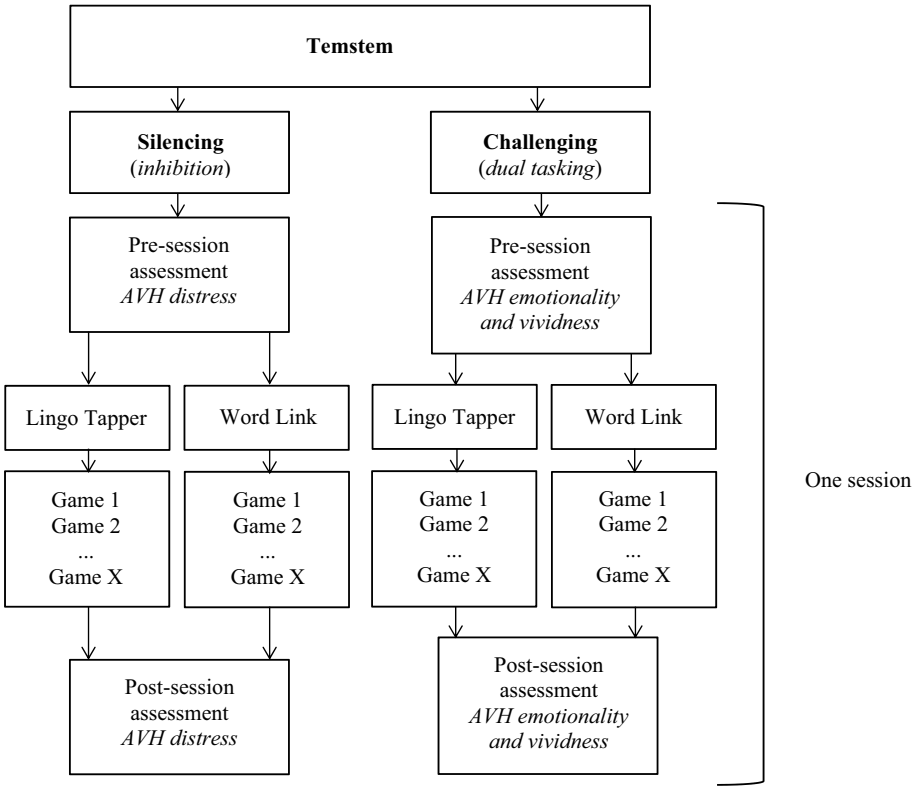


Fig. 1. Procedure in a Temstem session.

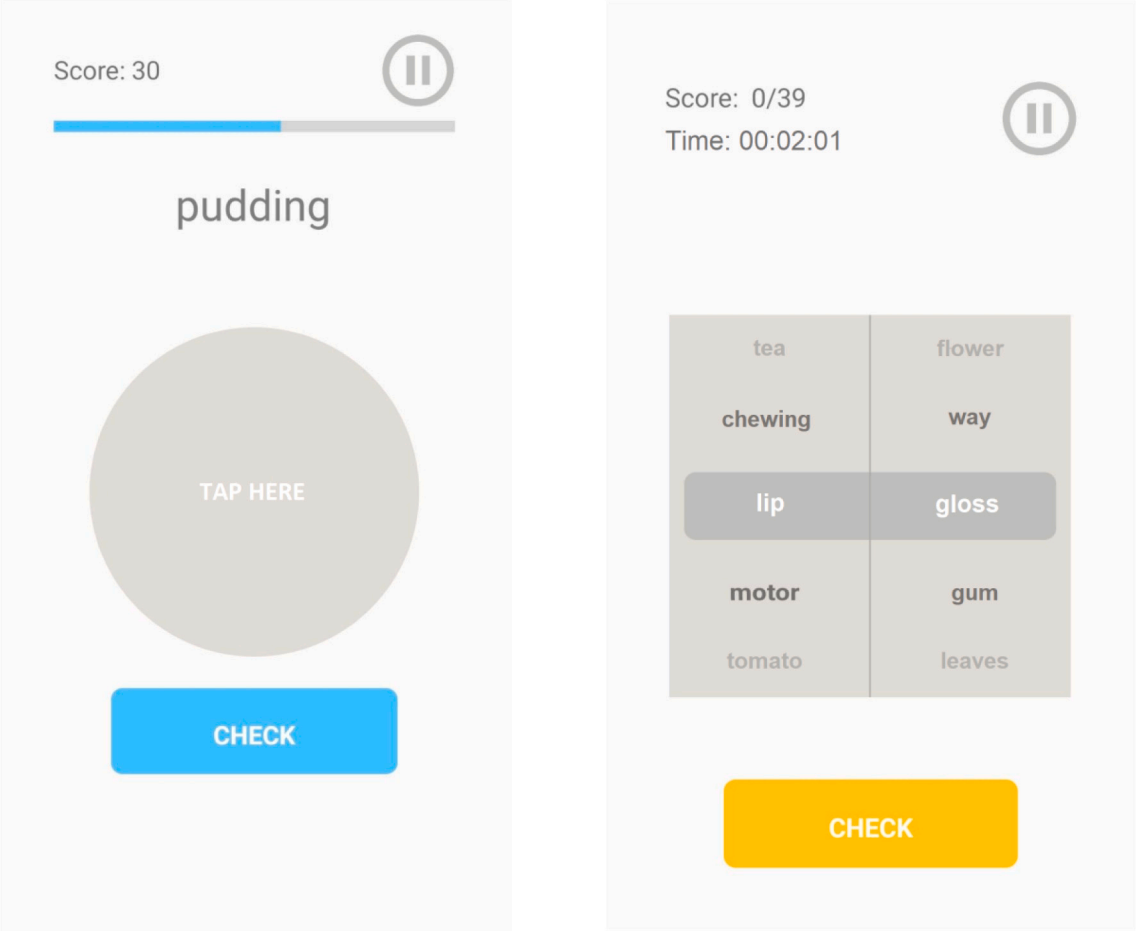


Fig. 2. Lingo Tapper (left) and Word Link (right).

1d) Is there a change in baseline voice-hearing distress (distress scores at the start of each new session) after multiple use?

Challenging (dual tasking)

2a) Is there a momentary effect of Challenging sessions on the emotionality and the vividness of voice memories?
2b) Are there differences between the Lingo Tapper and Word Link games in terms of momentary effects on the emotionality and the vividness of voice memories?
2c) Are effects on emotionality and vividness of voice memories sustained in multiple use?
2d) Are there changes in the baseline emotionality and vividness of voice memories after multiple use?

2. Method

2.1. Participants

According to Dutch law, the current research study was not subject to the Medical Research Involving Human Subjects Act, WMO (CCMO, 2020); medical ethics review was therefore not required. Informed consent was mandatory to permit collection and storage of user data. Data were collected between 14 February 2017 and 27 August 2019. Any person in the Netherlands was allowed to download and use Temstem freely. Temstem is disseminated in Dutch networks of mental health professionals and voice-hearing individuals, and it is advertised on multiple websites.

Users were included in the sample who (1) gave informed consent for the use of their anonymised data for research and evaluation purposes; (2) reported hearing voices; (3) completed at least level 2 (out of 6) in one of the games and/or had used Temstem at least 15 times (regardless of the functions or games used). These criteria enabled us to identify and process data from serious users who employed Temstem sufficiently to enable detection of effects. A further inclusion criterion was (4) an age from 16 to 89; users above 89 were excluded, due to a presumed greater chance of errors in the self-reported data. Users who reported hearing AVHs longer than their self-reported age were also excluded.

We also excluded certain game sessions recorded by participating users. For instance, we observed that the registration of the time the app was used, continued when the user clicked the app away (instead of using the 'stop' button); this resulted in time registrations of multiple hours. To overcome this, we removed any sessions longer than three standard deviations above the mean duration. We also examined how many times each participant used Silencing and Challenging respectively and removed data on the last 10 % of games played by the highest-frequency users, in order to avoid disproportionate influence from that small segment of users.

2.2. Procedure

Temstem is available on Android and iOS. After downloading Temstem, each user is asked to provide informed consent to share their data with researchers; decliners may continue using the application. Users then report a limited set of demographics, including age, and are introduced to Tim, the avatar. Tim provides basic psychoeducation on AVHs and explains the rationale of the app's functions. The user can re-read this information at any time. After the introduction, the use of

Temstem can begin.

A Temstem session includes the following steps (Fig. 1). The user selects the Silencing or Challenging function. If Silencing is selected, Tim asks the user to rate the current level of momentary AVH distress on a 1-to-7 Likert scale. Then the user selects which game to play: Lingo Tapper (LT) or Word Link (WL; see Fig. 2). In Lingo Tapper, successive words appear on the screen and the user is to tap the screen as many times as there are syllables in the word, doing so as fast as possible; thus the word *superman* would get three taps. Word Link is a language game showing two wheels of words, whereby the user is to correctly link two words together to form one new word. In Dutch, combined words are written as one word, for example *appel* and *boom* form *appelboom* (apple tree) and *thee* and *blaadjes* form *theeblaadjes* (tea leaves).

Users can play as many games as they want for as long as they want. Once the session is finished, a user is asked to re-rate the level of current AVH distress. In the event of a neutral or positive difference in pre- and post-game scores, the user receives visual feedback on that difference: the pre- and post-game scores are shown with the message: well done! The session is then closed.

In Challenging, the user follows the same procedure as with Silencing with a few exceptions. After selecting Challenging, the user first recollects a recent voice statement for 10 s, so as to activate that statement in the user's working memory; this is in line with the procedure in the dual-tasking lab studies cited above (Jongeneel et al., 2020; Matthijssen et al., 2019). The user then rates both the emotionality and the vividness of that memory on 7-point scales and then selects the cognitive-affective response that best describes what they think or feel when hearing the voice's content (neutral, weak, powerless, unsafe, inferior, bad, guilty). The user chooses between LT or WL and the game is started.

While playing, Tim additionally provides feedback based on the previously indicated cognitive-affective domain (for instance, if a user has tapped 'powerless', Tim says, 'You are strong!'). After the game, the user re-rates the emotionality and vividness of the memory.

For Lingo Tapper, the user is rewarded with a 'star' when he improves his personal score twice or when he plays the game several times. For Word Link, a star is received when a third of the total word combinations are found. When three stars are acquired, the user reaches the next level of that game.

2.3. Measures

As described, AVH distress was rated before and after the use of the Silencing function, and emotionality and vividness of voice memories were rated before and after the use of Challenging. All variables were measured on 1–7 Likert scales, with a score of 1 representing no distress, emotionality or vividness at all and 7 indicating maximum distress, emotionality or vividness. Baseline distress, emotionality and vividness levels refer to the corresponding ratings before the start of each session. Mean change scores for distress, emotionality and vividness were calculated by first totalling change scores per session and then calculating the mean change across sessions for each user.

2.4. Statistical analysis

Data were analysed using SPSS Statistics, version 23. First, to examine differences in pre- and postscores of voice-hearing distress after a Silencing session (research question 1a), we employed a two-tailed paired-samples *t*-test with user's mean baseline (pre-session) distress scores and mean post-session distress scores as variables. Using the average of the distress measures for each user allowed us to take the within-subject clustering into account without violating the assumption of independence. Next, to test whether differences between pre- and post-distress scores differed between the two games (question 1b), a two-tailed paired-samples *t*-test was computed with the mean changes in distress for Lingo Tapper and for Word Link as variables. Lastly, two linear mixed models with random intercepts, fixed slopes, and variance components structure were computed to test whether the number of sessions performed (independent variable) predicted the change in distress (1c) and the baseline distress score (1d) (dependent variables). For these analyses, we included all sessions; in order to take user-level clustering into account, the analyses consisted of two levels: sessions (level 1) clustered within individuals (level 2). Subsequently, the same three-step analysis approach was carried out for emotionality and vividness of voice memory scores after the use of Challenging (2a–2d).

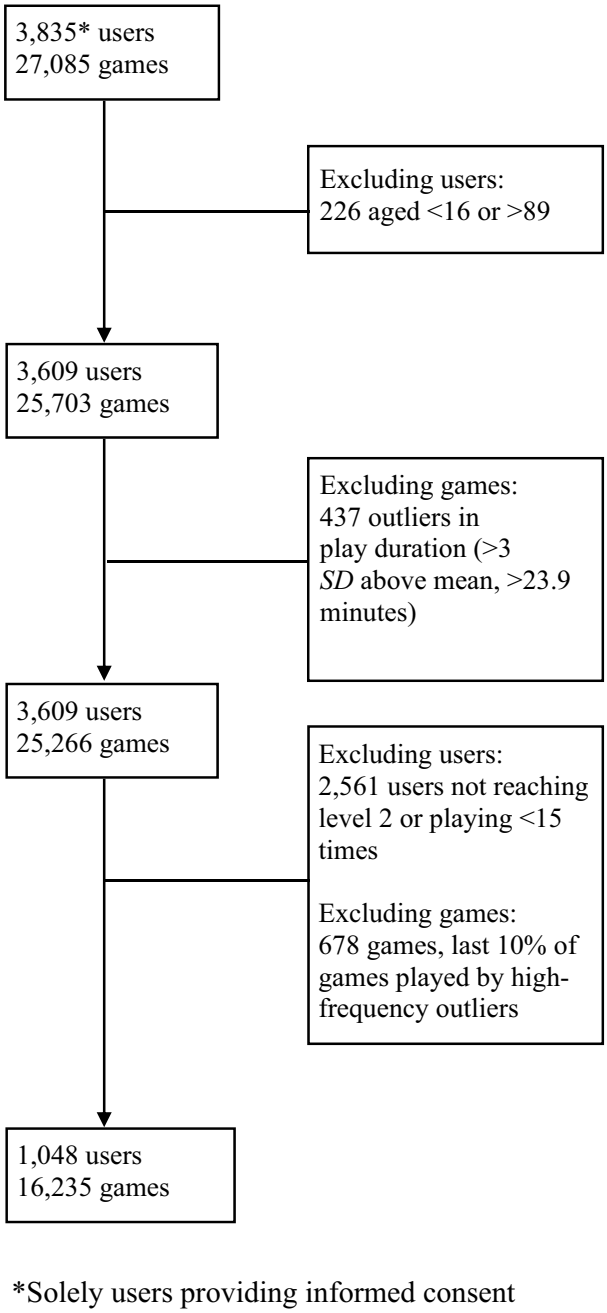


Fig. 3. Flowchart of the sample selection process.

2.4.1. Sensitivity analysis

As mentioned in Section 2.1 Participants, we removed the data of the last 10 % of games played by the highest-frequency users (outliers) in order to avoid potential disproportionate influence from these users. We added extra sensitivity analysis whereby this data was included, to test whether they indeed influenced the results.

Table 1
Momentary effects of inhibition on AVH distress and of dual tasking on the emotionality and vividness of voice memories.

	t	*df*	Pre-session score		Post-session score		Difference			*p*	Cohen's *d*
			M	*SD*	*M*	*SD*	*M*	*SD*	95 % CI		
Distress	27.12	958	5.18	1.52	4.37	1.77	0.81	0.93	0.76–0.87	<.001	0.49
Emotionality	23.16	651	5.73	1.30	4.60	1.70	1.13	1.25	1.04–1.23	<.001	0.74
Vividness	22.20	651	5.59	1.31	4.52	1.64	1.06	1.23	0.97–1.16	<.001	0.71

Table 2
Mean differences between Lingo Tapper and Word Link games in terms of momentary AVH distress and the emotionality and vividness of voice memories.

	t	*df*	LT		WL		Difference			*p*	Cohen's *d*
			M	*SD*	*M*	*SD*	*M*	*SD*	95 % CI		
Distress	1.69	594	−0.84	0.98	−0.91	1.05	0.08	1.08	−0.01–0.16	.092	0.07
Emotionality	2.88	226	−1.08	1.17	−1.36	1.49	0.28	1.48	0.09–0.48	.004	0.21
Vividness	2.29	226	−1.02	1.18	−1.24	1.43	0.22	1.47	0.03–0.42	.023	0.17

LT = Lingo Tapper; WL = Word Link.

3. Results

A total of 3835 people, who played games 27,085 times, gave informed consent to use their data for research and evaluation purposes (see Fig. 3 for the flowchart). After excluding users or games that did not fulfil inclusion criteria or that met exclusion criteria, and excluding outliers, our final sample included data from 1048 individuals who used Temstem to some extent and completed 16,235 games.

Of those 1048 participants, 79.0 % reported having been informed of Temstem by their mental health professional or coach, indicating that most users had received some form of mental health care. A further 8.8 % learned of Temstem from family, friends, colleagues or acquaintances, 7.7 % from the internet (e.g. a search engine), 1.2 % from a leaflet, poster or other media, and 3.2 % from other sources. Some 75.3 % of the users identified as female, 22.8 % as male, and 1.8 % reported neither gender. Ages ranged from 16 to 77 years ($M = 35.34$, $SD = 14.03$). On average, participants had been hearing voices for 4.95 years, ranging from 0 to 58 years ($SD = 8.74$). The total number of Temstem sessions per user ranged from 2 to 110 (total 16,235 games, $M = 15.49$, $SD = 16.88$). The number of Silencing sessions ranged from 1 to 76 (total 12,318 games, $M = 15.54$, $SD = 16.90$) and the number of Challenging sessions from 1 to 34 (total = 3917 games, $M = 7.35$, $SD = 7.62$). The Lingo Tapper game was played 10,064 times ($M = 10.31$, $SD = 13.95$) and Word Link 6171 times ($M = 6.69$, $SD = 7.61$).

3.1. Silencing

On average, momentary AVH distress was significantly decreased after a Silencing session, with a medium effect size, $t(958) = 27.12$, $p < .001$, $d = 0.49$ (Table 1). There was no significant difference in effectiveness between the Lingo Tapper and Word Link games, $t(594) = 1.69$, $p = .09$ (Table 2). The degree of distress reduction per session did not change significantly with more frequent use, $F(1, 7905.57) = 1.91$, $p = .17$, indicating that the reductive effect on distress remained stable in multiple use (see Fig. 4). The distress scores that users reported before the start of each session (baseline distress) did not change significantly in repeated use, $F(1, 11{,}741.38) = 0.78$, $p = .38$.

3.2. Challenging

Both the momentary emotionality of voice memories, $t(651) = 23.16$, $p < .001$, $d = 0.74$, and the vividness of the memories, $t(651) = 22.20$, $p < .001$, $d = 0.71$, decreased significantly after Challenging sessions, with medium to large effect sizes for both attributes (Table 1). The Word Link game showed significantly larger decreases in emotionality, $t(226) = 2.88$, $p < .005$, $d = 0.21$, and vividness, $t(226) = 2.29$, $p < .05$, $d = 0.17$, as compared with Lingo Tapper (Table 2). The degrees of change in emotionality and vividness during Challenging sessions diminished slightly with frequent use, $F(1, 2222.86) = 7.21$, $p < .05$; $F(1, 2289.92) = 4.25$, $p < .05$; for each time played, the effects on emotionality and vividness declined by 0.014 (1.3 %) and 0.010 (0.9 %) respectively. Baseline emotionality and vividness scores (at the start of each session) did not change significantly over multiple sessions, $F(1, 3705.96) = 3.11$, $p = .08$, $F(1, 3668.65) = 1.41$, $p = .24$ (Fig. 4).

3.3. Sensitivity analysis

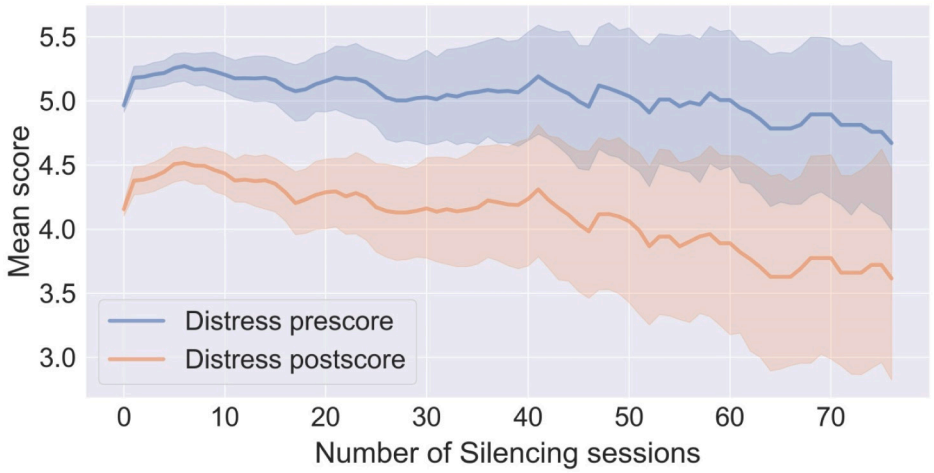
For sensitivity analysis, we included in addition to our primary dataset the last 10 % of games played by the highest-frequency users. This dataset contained data of 1048 individuals who played 18,118 games. Results are shown in Appendix A. Results of hypotheses 1a – 2b are in line with our above described results. However, a few results are not in line with our primary analyses.

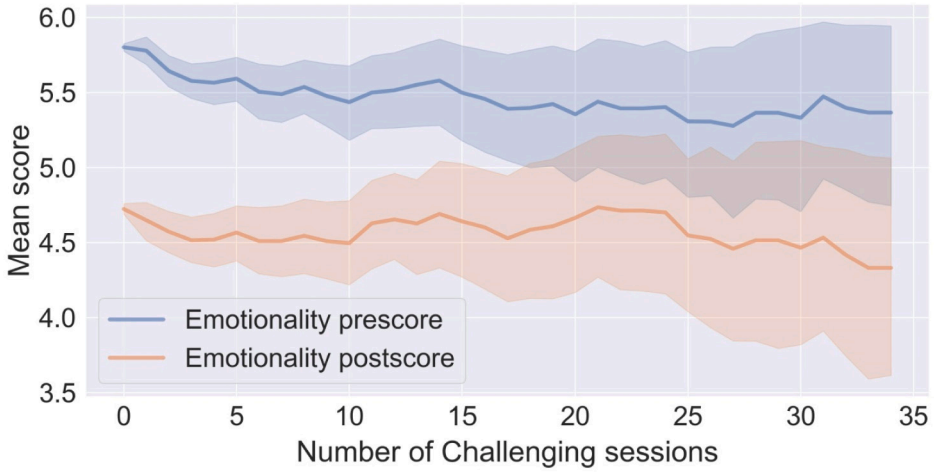
In this extended dataset, we found a significant reduction of the effect of Silencing on distress over time, $F(1, 8806.788) = 4.08$, $p < .05$. In addition, baseline distress decreased over frequent use, $F(1, 12{,}994.449) = 19.65$, $p < .001$. The effects of frequent use on the reductions of emotionality and vividness, as the effect of frequent use on baseline emotionality and vividness remained non-significant.

4. Discussion

This study, based on analyses of naturalistic user data, investigated whether momentary voice-hearing distress as well as the emotionality and vividness of voice memories were reduced after a Silencing or Challenging session in the Temstem mobile mental health app. We observed, in data of 1048 users, that momentary voice-hearing distress was significantly decreased after a Silencing session through both types of language games provided. This is in line with previous research showing that auditory, subvocal techniques such as humming and gargling are effective in inhibiting voices (Bick and Kinsbourne, 1987; Erickson and Gustafson, 1968). In addition, we found that the degree of reduction in momentary voice-hearing distress remained stable over multiple-session use; though that finding reflects multiple use only, it seems to indicate a sustainable coping effect for voice-hearing individuals. Although ‘coping is not cure’, an important advantage of Temstem is that it can be applied in many situations, such as travelling on public transport or being at home or work.

After a Challenging sessions, voice memories were less emotional and vivid, at least momentarily. This is in line with results from our previous lab study on this topic (Jongeneel et al., 2020), which showed that a dual task – thinking about a voice memory while using the Lingo





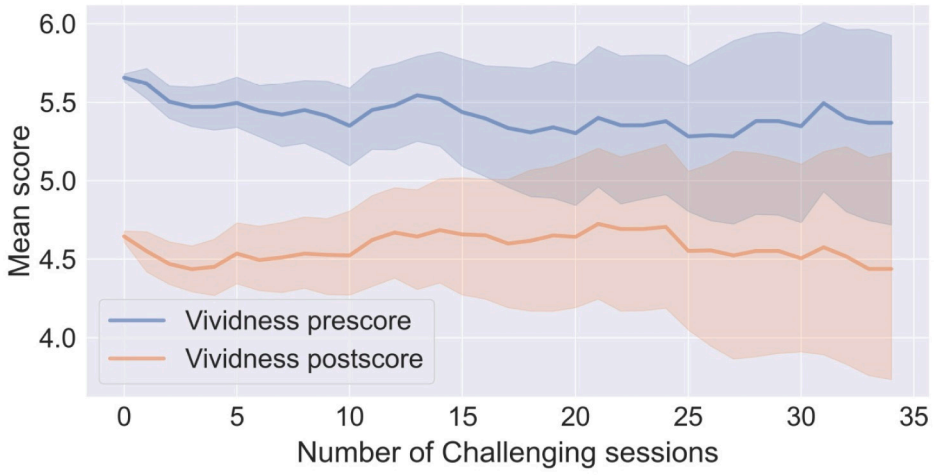


Fig. 4. Baseline momentary distress, emotionality and vividness during repeated use of Temstem.

Tapper game – decreased the emotionality and vividness of the voice memory in comparison with recall only. Interestingly, the current study found that reductions in emotionality and vividness in voice memories were larger after a Word Link game than a Lingo Tapper game. Since some research suggests that higher-taxation tasks result in greater reductions in the vividness and emotionality of images (e.g. Van Schie et al., 2016; van Veen et al., 2015), this may indicate that Word Link is a more cognitively taxing game than Lingo Tapper. Whether effects might generalise to a reduction of AVH distress in daily life, should be investigated further.

Furthermore, after using Challenging multiple times, reductions of emotionality and vividness of voice memories diminished slightly; for every game played, the reduction of emotionality and vividness declined by 1.3 % and 0.9 % respectively. Those small percentages do not appear very relevant from a clinical perspective, since participants had to use a language game approximately 39 times before the effect on emotionality

was halved. New Temstem tasks could probably be developed in future to minimise that process.

71 % of the persons who downloaded Temstem did not or very limitedly use Temstem (they did not reach level 2 of one of the games or used Temstem below 15 sessions), indicating that they were not or only minimally exposed to the Temstem functionalities, and were therefore not included in the analyses. Above described results only account for the users who were exposed to the functionalities. Temstem is evidently based on a 'use it or lose it' principle and it is of importance to acknowledge that Temstem is not effective for a part of the persons who download the app.

Results of the sensitivity analysis, including outliers, were largely in line with our primary results. However, there were two contradictory results in comparison with our primary analysis. For instance, we did not initially find a decrease in baseline distress over frequent use, but in the dataset containing outliers we did find a significant decrease. This demonstrated that a few high-frequent users can have a great influence on results.

4.1. Strengths and limitations

We believe this study has several strengths. In this study, we explored self-reported momentary user data of an eHealth app (other than log data or data collected by experience sampling method apps). Although such naturalistic data is subject to biases and confounding when participants are not randomised and the causes of observed effects remain uncertain (Fagiolini et al., 2017), naturalistic real-world studies of apps include data from all serious users. As a result, outcomes may be generalisable to a heterogeneous population of other users who were exposed to the Temstem functionalities, making this a suitable study design for evaluating the effects of mobile health (mHealth) interventions (Fagiolini et al., 2017). The current study included voice-hearing individuals of varied age, voice-hearing duration, degree of AVH distress, mental health care involvement and other differentiating characteristics; data from over 1000 frequent users were analysed.

There are also several important limitations to our study. First, we did not have a control arm in this naturalistic study; therefore, we cannot rule out that e.g. voice-hearing distress would also have been decreased without Temstem. Also, the user choose autonomously what function and what game to play, this was not randomised. Because of these study characteristics, results should be interpreted with caution. Second, results are based on self-reports; we cannot rule out effects of social desirability on reported answers. Third, some users might have received some professional support in using Temstem. Unfortunately, we have no insight into how many users received support and to what extent they were guided. That would be important to know, as a recent meta-analysis showed that *stand-alone* use of an app usually yields no or a very small effect (Weisel et al., 2019). Its authors therefore concluded that a professional should monitor the progress and guide the user (in a blended approach) so as to possibly enhance effects. Guidance improves adherence and lowers dropout rates, and the intervention can be tailored to the individual's needs (Andersson and Titov, 2014). Hence, the momentary effects from Temstem may increase if it is provided in a blended fashion, but to whom and to what extent that might apply should and will be investigated further. Fourth, context variables were not measured. As a consequence, (1) we had limited insight into the population characteristics and were not able to identify which individuals with what characteristics did or did not profit from the momentary Temstem effects; (2) we do not know to what extent users were experiencing distressing voices at other points in time when they were not using Temstem; and (3) the reasons why users continued or stopped using Temstem remain unclear. The lack of insight into context variables, might have implications for the interpretation of the long-term effects. For example, we do not know whether persons stopped using Temstem because the effect of Temstem on voice-hearing distress or emotionality or vividness of voice memories decreased over time. In this case, the stable long-term effect of Temstem we reported is based on a sample of persons who do not experience such a decrease and the long-term effect is actual an over-estimation of the reality. We have to learn more about what individuals download Temstem, who uses it for a short while or a longer period, who stops using it, and what a person needs to benefit from Temstem. In short, our results, and in particular any indications of longer-term effects, should be interpreted with caution.

4.2. Conclusions

In this heterogeneous population of voice-hearing individuals with varying levels of AVH distress, Temstem might be a promising momentary coping tool. After a Silencing session, voice-hearing distress was momentarily reduced and after the dual-tasking function emotionality and vividness of voice memories were momentarily reduced. We are currently analysing results from a randomised controlled trial to study the effects of Temstem as a stand-alone intervention in a more homogeneous population of individuals with severe voice hearing. We have learned several important things. First, assessing the context surrounding the use of a mental health care app is essential, and suitable data collection should be included in every app that is analysed for effectiveness. Second, complementary analysis of both naturalistic and RCT data should be performed more often to further validate findings concerning effectiveness in the general population. Temstem is a promising tool that we will continue to adapt and evaluate together with users and professionals to further enhance its effectiveness.

Funding

The authors thank the management of the Parnassia Groep (Monsterseweg 93, 2553 RJ Den Haag, Netherlands) for funding the development and maintenance of the Temstem app and for providing financial support to conduct this research. The Parnassia Groep had no involvement in this study.

Declaration of competing interest

None.

Acknowledgement

We thank all users who allowed us to analyse their data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.invent.2022.100580.

References

Anderson, R.M., Funnell, M.M., 2005. Patient empowerment: reflections on the challenge of fostering the adoption of a new paradigm. Patient Educ. Couns. 57, 153–157. https://doi.org/10.1016/j.pec.2004.05.008.

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

Bauer, S.M., Schanda, H., Karakula, H., Olajossy-Hilkesberger, L., Rudaleviciene, P., Okribelashvili, N., Stompe, T., 2011. Culture and the prevalence of hallucinations in schizophrenia. Compr. Psychiatry 52 (3), 319–325. https://doi.org/10.1016/j.comppsych.2010.06.008.

Ben-Zeev, D., 2012. Mobile technologies in the study, assessment, and treatment of schizophrenia. Schizophr. Bull. 38 (3), 384–385. https://doi.org/10.1093/schbul/sbr179.

Bick, P.A., Kinsbourne, M., 1987. Auditory hallucinations and subvocal speech in schizophrenic patients. Am. J. Psychiatry 144 (2), 222–225. https://doi.org/10.1176/ajp.144.2.222.

Calabretta, N., 2002. Consumer-driven, patient-centered health care in the age of electronic information. J. Med. Libr. Assoc. 90 (1), 32–37.

CCMO, 2020. Retrieved from Ministerie van Volksgezondheid, Welzijn en Sport website. https://www.ccmo.nl/onderzoekers/wet-en-regelgeving-voor-medisch-wetenschappelijk-onderzoek/uw-onderzoek-wmo-plichtig-of-niet.

Cuijpers, P., Kleiboer, A., Karyotaki, E., Riper, H., 2017. Internet and mobile interventions for depression: opportunities and challenges. Depress. Anxiety 34 (7), 596–602. https://doi.org/10.1002/da.22641.

Daalman, K., Boks, M.P.M., Diederen, K.M.J., De Weijer, A.D., Blom, J.D., Kahn, R.S., Sommer, I.E.C., 2011. The same or different? A phenomenological comparison of auditory verbal hallucinations in healthy and psychotic individuals. J. Clin. Psychiatry 72 (3), 320–325. https://doi.org/10.4088/JCP.09m05797yel.

Erickson, G.D., Gustafson, G.J., 1968. Controlling auditory hallucinations. Psychiatr. Serv. 19, 327–329. https://doi.org/10.1176/ps.19.10.327.

Fagiolini, A., Rocca, P., De Giorgi, S., Spina, E., Amodeo, G., Amore, M., 2017. Clinical trial methodology to assess the efficacy/effectiveness of long-acting antipsychotics: randomized controlled trials vs naturalistic studies. Psychiatry Res. 247, 257–264. https://doi.org/10.1016/j.psychres.2016.11.044.

Farhall, J., Greenwood, K.M., Jackson, H.J., 2007. Coping with hallucinated voices in schizophrenia: a review of self-initiated strategies and therapeutic interventions. Clin. Psychol. Rev. 27 (4), 476–493. https://doi.org/10.1016/j.cpr.2006.12.002.

Gallagher, A.G., Dinan, T.G., Baker, L.J., 1994a. The effects of varying auditory input on schizophrenic hallucinations: a replication. Br. J. Med. Psychol. 67 (1), 67–75.

Gallagher, A.G., Dinan, T.G., Baker, L.J.V., 1994b. The effects of varying auditory input on schizophrenic hallucinations. Br. J. Med. Psychol. 67, 67–75.

Gallagher, A.G., Dinan, T.G., Baker, L.V., 1995. The effects of varying information content and speaking aloud on auditory hallucinations. Br. J. Med. Psychol. 68 (2), 143–155.

Green, M.F., Kinsbourne, M., 1989. Auditory hallucinations in schizophrenia: does humming help? Biol. Psychiatry 25 (5), 633–635. https://doi.org/10.1016/0006-3223(89)90225-4.

Huhn, M., Nikolakopoulou, A., Schneider-Thoma, J., Krause, M., Samara, M., Peter, N., Leucht, S., 2019. Comparative efficacy and tolerability of 32 oral antipsychotics for the acute treatment of adults with multi-episode schizophrenia: a systematic review and network meta-analysis. Lancet 394, 939–951. https://doi.org/10.1016/S0140-6736(19)31135-3.

Johns, L.C., Kompus, K., Connell, M., Humpston, C., Lincoln, T.M., Longden, E., Laroi, F., 2014. Auditory verbal hallucinations in persons with and without a need for care. Schizophr. Bull. 40 (4), 255–264. https://doi.org/10.1093/schbul/sbu005.

Jongeneel, A., Scheffers, D., Tromp, N., Nuij, C., Delespaul, P., Riper, H., Van Den Berg, D., 2018. Reducing distress and improving social functioning in daily life in people with auditory verbal hallucinations: study protocol for the temstem' randomised controlled trial. BMJ Open 8 (3). https://doi.org/10.1136/bmjopen-2017-020537.

Jongeneel, A., van Veen, S.C., Scheffers, D., Riper, H., van den Hout, M.A., van der Gaag, M., van den Berg, D., 2020. Linguistic dual tasking reduces emotionality, vividness and credibility of voice memories in voice-hearing individuals: results from a controlled trial. Schizophr. Res. https://doi.org/10.1016/j.schres.2019.11.048.

Maijer, K., Begemann, M.J.H., Palmen, S.J.M.C., Leucht, S., Sommer, I.E.C., 2018. Auditory hallucinations across the lifespan : a systematic review and meta-analysis. Psychol. Med. 48 (6), 879–888. https://doi.org/10.1017/S0033291717002367.

Matthijssen, S.J.M.A., Heitland, I., Verhoeven, L.C.M., van den Hout, M.A., 2019. Reducing the emotionality of auditory hallucination memories in patients suffering from auditory hallucinations. Front Psychiatry 10, 1–12. https://doi.org/10.3389/fpsyt.2019.00637.

Sommer, I.E.C., Slotema, C.W., Daskalakis, Z.J., Derks, E.M., Blom, J.D., Van Der Gaag, M., 2012. The treatment of hallucinations in schizophrenia spectrum disorders. Schizophr. Bull. 38 (4), 704–714. https://doi.org/10.1093/schbul/sbs034.

Thomas, N., Bless, J.J., Alderson-Day, B., Bell, I.H., Cella, M., Craig, T., Jardri, R., 2019. Potential applications of digital technology in assessment, treatment, and self-help for hallucinations. Schizophr. Bull. 45 (1), S24–S31. https://doi.org/10.1093/schbul/sby103.

Turner, D.T., Burger, S., Smit, F., Valmaggia, L.R., van der Gaag, M., 2020. What constitutes sufficient evidence for case formulation–driven CBT for psychosis? Cumulative meta-analysis of the effect on hallucinations and delusions. Schizophr. Bull. https://doi.org/10.1093/schbul/sbaa045.

Van Schie, K., Van Veen, S.C., Engelhard, I.M., Klugkist, I., Van den Hout, M.A., 2016. Blurring emotional memories using eye movements: individual differences and speed of eye movements. Eur. J. Psychotraumatol. 7 (1), 29476. https://doi.org/10.3402/ejpt.v7.29476.

Van Veen, S.C., Van Schie, K., Wijngaards-de Meij, L.D.N.V., Littel, M., Engelhard, I.M., Van den Hout, M.A., 2015. Speed matters: relationship between speed of eye movements and modification of aversive autobiographical memories. Front Psychiatry 6 (April), 45. https://doi.org/10.3389/fpsyt.2015.00045.

Weisel, K.K., Fuhrmann, L.M., Berking, M., Baumeister, H., Cuijpers, P., Ebert, D.D., 2019. Standalone smartphone apps for mental health—a systematic review and meta-analysis. Npj Digit. Med. 2 (1), 1–10. https://doi.org/10.1038/s41746-019-0188-8.