

Cognitive-Behavioural Therapy in Personal Devices: A Case Study

Ray Pohjanheimo, MSc.

Faculty of Electrical Engineering, Mathematics and Computer Science

University of Twente, Enschede

29/10/2018

### **Abstract**

A case study was performed to test self-reported emotional valence improvement and usability of a personal device application “Free CBT”. Free CBT was developed based on earlier therapeutic framework of cognitive-behavioural therapy with the aim to improve mental well-being. The hypotheses were that the appraisal for self-reported emotional valence is increased from before to after using the application and that the application is rated above neutral in usability. The results of an incident diary study and a system usability scale showed significant support for both self-reported emotional valence improvement and usability. Implications, limitations, and future suggestions are discussed.

*Keywords:* cognitive-behavioural therapy, personal device, self-help, human-computer interaction, emotional valence

## Cognitive-Behavioural Therapy in Personal Devices: A Case Study

Cognitive-Behavioural Therapy (CBT) is a therapeutic framework within psychology, with first instances as early as in the 1950s and 1960s (Ellis, 1957; Beck, 1960). It focuses on changing an individual's cognition and behaviour through a variety of different techniques established by a plethora of researchers in the field. Cognition refers to an individual's ability to direct their thoughts and actions (Miller, 2018). CBT has shown positive results in improving mental health as shown by a meta-analysis by Butler, Chapman, Forman, and Beck (2006), making it an effective tool for therapeutic purposes. CBT has been particularly potent when applied to the treatment of several anxiety and depressive disorders, such as generalized anxiety disorder and unipolar depression. Curiously, CBT was shown to be more effective in treating adult depression than antidepressants (Butler et al., 2006). The results of the meta-analysis show an overall pattern for multiple therapeutic applications.

However, a lot of people don't have access to CBT or other mental health treatments, as illustrated by the whopping estimated 83 million EU citizens suffering from a mental health issue ("Data and resources", 2018). The amount of available therapists per population varies drastically within the EU, with some countries lacking in resources they can provide for people in need of mental health care. The use of CBT is mostly only applied in therapeutic environments, however, some of the techniques rely on an individual's own effort without supervision. Therefore, parts of CBT could also be helpful for people who are not enrolled in therapy. The problem the present research

seeks to tackle is could an easily accessible and easy to use CBT-based application on a personal device alleviate issues to an extent or simply just improve the user's mood.

CBT's utility in mobile applications has been explored by previous research. A study by Koffel et al. (2016) investigated if a CBT application was an impactful supplement in treating insomnia for patients currently taking part in cognitive-behavioural therapy. The results of the study showed that the app did not interfere with therapy and that the users were content with using the application. However, the study compared two groups who were both enrolled in the therapy and since both groups improved in their treatment outcomes, no conclusions could be drawn about the app's effectiveness in treating patients outside therapy. Research in non-therapeutic settings could therefore provide important insights to the applicability of CBT in personal devices.

Another study within the topic studied the effectiveness of a CBT application in decreasing negative moods and increasing positive moods. A study by Kinderman et al. (2016) investigated the impact of a CBT based application that prompted users to assess an upsetting situation as if someone else had made the appraisal they originally made. The results showed a statistically significant decrease in negative mood intensity as well as an increase in positive mood intensity, which is a promising basis for the present research. While the application used in the research by Kinderman et al. (2016) had a prompt for users to re-evaluate their appraisal, a specific identification process for cognitive biases was still missing.

One of the techniques of CBT helps people identify cognitive biases that might have affected their appraisal of an incident or thought process, which is what this study focuses on. The technique was chosen because of its relative ease to be transformed into an application design that an individual could use on their own without intervention. An application by the name of “Free CBT” was developed based on this technique. The purpose of Free CBT is to provide an accessible way for people to apply the technique in their everyday life to improve their general well-being. As such, the application does not give diagnostic suggestions, nor does it explore any other features that would require the advice and intervention of a therapist. The research questions I sought to answer were if the chosen methods provide insights and results, if the app improves self-reported well-being, and what is the usability of the application. The hypotheses set were as follows:

*H1: Using Free CBT improves self-reported emotional valence.*

*H2: The usability of Free CBT is rated above neutral.*

The hypotheses are studied in a one-participant case study combining multiple different user research methods. This is to pilot the initial concept in a small and manageable manner in favour of a large sample study, as well as to test the effectiveness of various experimental methods in gaining knowledge about how the effectiveness and usability of the application can be measured. As such, the results of this research aim to give initial insight into the application’s value but not necessarily draw defining conclusions.

## **Application Design**

The application design followed a simple five-screen format. The application started with a short prompt asking “What’s on your mind?” (Figure 1). The application further said “Do you feel bad about something? What was it? You can either type about it in the box or just think about the incident.”, The user was given an opportunity to elaborate on the event in a text box, however, the user was informed that the elaboration is optional. A disclaimer at the bottom of the page stated that user text input was not recorded. The second screen (Figure 2) asked the user to rate the amount of distress caused by the event by using a slider ranging from “Not at all” to “Very distressing”. The third screen (Figure 3) presented the cognitive biases with short descriptions, adapted from the work of Beck (1995). The user was instructed to tick the checkboxes of any and all applicable ones that might have affected their appraisal of the situation. The fourth screen was largely identical to the second screen (Figure 2) and prompted the user to re-evaluate their level of distress after considering the cognitive biases. The last screen simply represented the end of the flow and wished the user would feel better, with a button redirecting to the start page if wanted.

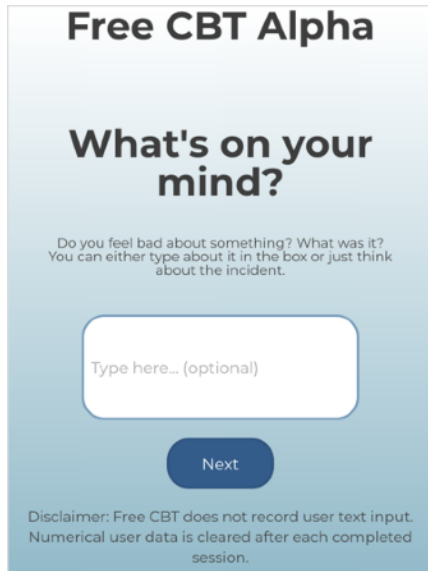


Figure 1. Start Screen

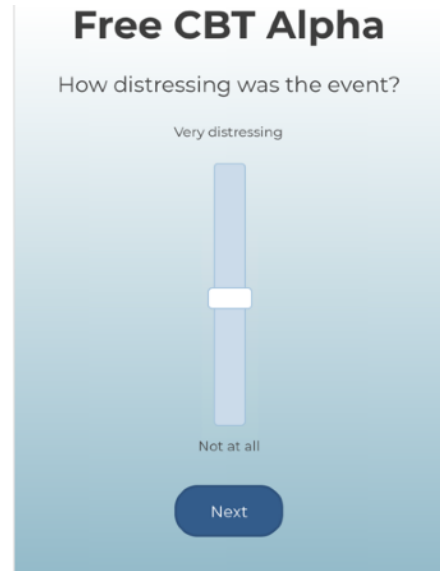


Figure 2. Distress Scale



Figure 3. Cognitive Biases

## Method

### Participants

One participant from the University of Twente volunteered to take part in the experiment. While there was no compensation offered, the nature of the programme encouraged voluntary participation as part of the learning experience and to help other students achieve their research goals. The participant's age and gender were not recorded to ensure anonymity and because they were not the factors of interest for the purposes of this research, however, it can be declared that the participant was legally an adult. An initial briefing revealed that the participant was well versed with technology, however, this was not quantified. The participant was asked to take part in all the phases of the experiment.

### Materials and Procedure

Incident diary study, semi-structured interview, field study, and methodological assessment were chosen as the instruments for the study. The decision for including the chosen instruments were based on recommendations in *Understanding your users* (Courage & Baxter, 2005). The diary study was chosen to gain an understanding to how the application was used in everyday life outside of laboratory settings. The function for the interview was to receive deep knowledge about what the participant thought of the application and what kind of features could be included in the future. For the field study, the aim was to see the participant in action with the application and observe if there were



any hiccups in the user flow. Methodological assessment was chosen for its function in quantitative data collection for usability measurement. All of the parts of the study were designed as such that there was no deception used, no personally identifiable information was gathered, and the participant was presented with informed consent of their voluntary participation prior to the experiment. The procedure of the experiment at large started with a three-week diary study phase, moving on to the interview after the data from the diaries were collected, and ending with a field study session followed by the methodological assessment. All the parts of the experiment took place either in a private residence in Enschede or out in the field for the case of the diary study.

*Diary Study.* The incident diary format was chosen because it was predicted that the application would not be used regularly but instead at specific instances. The diary contained three open questions about the usage incident and appraisals, two Likert-scales assessing self-reported emotional valence before and after use ranging from one to nine, and an optional prompt for additional comments (see Appendix A). The open questions were preceded with the statement “Describe in as much detail as you want” to not pressure the participant to overshare personal information they would not be comfortable sharing. Participant was provided with a link to the application along with 20 copies of the diary study template and they were instructed to fill in an entry whenever they used the application over the period of three weeks.

*Interview.* The semi-structured interview consisted of a set of pre-made questions as well as spontaneous questions during the interview (Appendix B). The questions were directed towards user experience, the usefulness, privacy concerns, and approachability.

The interview was conducted in a private residence, recorded, and later transcribed with a software application called Descript for accurate analysis. Time allotted for the interview was 45 minutes.

*Field Study.* The field study was conducted as a deep hang out in a private residence. Deep hang out was chosen because the primary interest was a holistic view of the entire process. The participant was instructed to engage in thinking aloud while completing three runs of the application. No detailed questions or considerations were provided for the participant, as per the plan of the deep hang out.

*Evaluation Method.* The final part of the study was the evaluation method, in particular, the System Usability Scale (SUS) questionnaire by John Brooke (1996). The SUS was slightly adapted by increasing the Likert scale to seven in order to capture a more accurate scope of the usability (see Appendix C). The SUS was presented after the field study.

## **Results**

*Diary Study.* The collected data from the diary study showed four instances where the participant had filled in the information. The incidents where the participant used the app were varying and showed no distinct pattern. The diary study's open questions gave an impression that the participant's mood was generally improved after using the app. The statements for what was the participant's appraisal after using the app were all deemed positive and varied from "Felt a little better for sure" to "Way better, definitely help me slow down and think, and analyze, and step back from the immediate emotions

of the situation.” The difference in mean emotional valence between before ( $M = 3.0$ ,  $SD = 1.4$ ) and after ( $M = 7.0$ ,  $SD = 1.4$ ) was noticeable, although due to the low sample size, no statistical testing was feasible.

*Interview.* The interview contained no quantifiable measures. However, it provided important insights to how privacy should be ensured and also how additional information about personal usage patterns could improve the experience. Answers such as “- - - I could imagine it would be really nice to have someone who was tracking your data who then - like if your therapist was watching your data - then you could meet with them every week or two and then go over it and see how you're improving.” showed that some additional features and integration to a therapeutic setting could be a beneficial future development. The interview also gave insights to monetising the application: “I mean, I probably preferred to pay up, I would probably prefer to pay upfront but realistically I would probably have ads. Because in this day and age 99 cents is just too much to ask. - - -” Lastly, the interview answered some questions about privacy, in statements such as “- - - You're sharing with your computer how you're feeling and so you kind of have to be comfortable with that. And once you're comfortable with that the privacy issue is a lot less but it's definitely on your mind. So you know, it would have been nice if there were - or I mean there is a disclaimer at the beginning, right? - - - does not record any user data no, and so that was kind of soothing, or reassuring, that's the right word.” The insights from the interview were particularly important for considering future developments in the application.

*Field Study.* The results of the field study were rather impoverished and did not add any unique insights. The participant completed the application run three times while thinking aloud but nothing remarkable came up during the experiment. While the field study did not yield any additional information, the fact that there were no obstacles that could have prompted some deeper elaboration could speak for a good user experience since the flow was smooth.

*Evaluation Method.* The SUS that the participant completed after the field study displayed a trend of high usability ratings across the items ( $M = 6.4$ ,  $SD = 0.7$ ). This shows initial support for the second prediction.

### **Discussion**

The first hypothesis, using Free CBT improves self-reported emotional valence, gained initial support based on the analysis of the diary study entries. The second hypothesis, the usability of Free CBT is rated above neutral, also received support based on the analysis of the SUS. However, since there was only one participant, statistical testing was not feasible. Nevertheless, the case study showed promising results in the capability of Free CBT in increasing well-being and being a usable application. The results have the potential to serve as a basis for researching and developing the application further.

The diary study was the most promising method because it showed some budding evidence for a trend across instances. Further, it gave insights to real world situations where the application might be used. The interview was effective in finding out deeper opinions and improvements that the other methods could not provide. The field study was

the least effective, possibly because the application was very straight forward and did not require a lot of thinking or elaboration. The usability scale was the most informative about the usability of the app but limited in other factors.

### **Implications**

The positive results of the case study could reflect an underlying pattern for Free CBT's use in increasing mood and potentially alleviating symptoms of mental health issues. The usability ratings could also prove to be generalisable in a wider sample, suggesting that the application would be easy to adopt by the population. Overall, the results show support for a potential solution that could help people outside of therapy or perhaps even as an assisting tool for therapeutic work. Regarding the methods, the results suggest that the diary study and the system usability scale were the best tools for measuring the application's effectiveness in quantitative terms.

For qualitative information, the interview was by far the most effective instrument. While the diary study provided some information about the context of use, motivation, and appraisals, the interview truly delved into the thoughts and opinions of the participant. Important considerations, such as privacy and additional features, were unraveled during the interview, and while the initial belief already was that for such a personal application privacy protection is essential, receiving confirmation from the participant increased the importance of the conviction. Regarding privacy, the participant had stated that collecting private information would be appropriate if the importance of data collection was highlighted and possibly assessed by a therapist. For additional features, personal progress and usage patterns were named as the most essential ones,

however, these are intertwined with user data collection. Therefore implementing the features would require careful assessment of how to inform the user of data collection and how to let the user opt out of the data collection.

### **Limitations**

Since the study only included one participant, the results are not very generalisable to a wider population yet. It is also noteworthy to point out that the participant was an avid user of technology, so the results might not be as applicable to non tech-savvy people. Further, the diary entries only contained four incidents which is a rather small sample even for a case study. Emotional valence was measured through self-reporting and while it was likely a fairly accurate representation of the participant's mood, physical measurements or recording user reports within the application could have provided a more accurate dataset. Nevertheless, a lot of insight was gained into what methods gave some substantial information about the user experience. Due to ethical considerations, the data collection was intentionally limited in personal and private information, which restricts the amount of insight in the user profile and the issues a user encounters. This however, is unlikely to change in future research unless the ethical considerations can somehow be controlled for and mitigated.

### **Conclusion**

Initial support for Free CBT's capability for improving mental well-being was provided by the research, as well as supporting evidence for the usability of the designed application. For future research, a larger sample size and fewer methods would be advisable in order to confirm whether there is a larger pattern of improvement in

emotional valence. In particular, the diary study and SUS were the most effective in gathering insights with quantifiable measurements. The other methods were useful - with the exception of field study - in collecting qualitative information, however, due to their labour intensity it would not be recommended to use interviewing or field studying for assessing how effective the application is for improving well-being. Besides self-reporting measures, emotional valence could also be measured through other methods, such as EEG, to gain even more convincing evidence of the phenomenon.

## References

- Ellis, A. (1957). Rational Psychotherapy and Individual Psychology. *Journal of Individual Psychology, 13*: 38-44.
- Beck, A. T. (1967). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.
- Beck, J. (1995) *Cognitive therapy: Basics and beyond*. New York: Guilford press
- Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry, 189(194)*, 4-7.
- Butler, A., Chapman, J., Forman, E., & Beck, A. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review, 26(1)*, 17-31.
- Data and resources. (2018). Retrieved from <http://www.euro.who.int/en/health-topics/noncommunicable-diseases/mental-health/data-and-resources>
- Koffel, E., Kuhn, E., Petsoulis, N., Erbes, C., Anders, S., & Hoffman, J. et al. (2016). A randomized controlled pilot study of CBT-I Coach: Feasibility, acceptability, and potential impact of a mobile phone application for patients in cognitive behavioral therapy for insomnia. *Health Informatics Journal, 24(1)*, 3-13. doi: 10.1177/1460458216656472
- MeasuringU: 10 Things to Know About the System Usability Scale (SUS). (2018). Retrieved from <https://measuringu.com/10-things-sus/>



Miller, E. (2018). Executive function and higher-order cognition: Definition and neural substrates. Retrieved from [http://www.academia.edu/2836377/Executive\\_function\\_and\\_higher-order\\_cognition\\_Definition\\_and\\_neural\\_substrates](http://www.academia.edu/2836377/Executive_function_and_higher-order_cognition_Definition_and_neural_substrates)

## Appendices

### Appendix A. Diary Study

*Free CBT Diary Study (<https://autocbt.bubbleapps.io/>)*

*ID:*

*Date:*

*Describe in as much detail as you want*

*What kind of a situation prompted you to use the app?*

*Did you encounter any difficulties in using the app? If so, please describe.*

*What was your appraisal after using the app?*

*On a scale from 1-9 how negative/positive did you feel before using the app?*

<i>Very Negative</i>			<i>Neutral</i>			<i>Very Positive</i>		
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>

*On a scale from 1-9 how negative/positive did you feel after using the app?*

<i>Very Negative</i>			<i>Neutral</i>			<i>Very Positive</i>		
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>

*Other comments:*

## Appendix B. Pre-made Interview Questions

### ***Warm-up questions***

*Are you familiar with self-help applications?*

*What do you know about the Cognitive-Behavioural Therapy?*

### ***Core***

*What are your thoughts about therapy applications?*

*How did you feel using the app?*

*If applicable, how does this app compare to therapy?*

*How often do you use technology in order to solve a personal issue?*

*What are your thoughts about using technology as a supplement to counselling or therapy?*

*If you had the choice between seeing ads on the application or paying a small upfront fee to use it, which one would you prefer?*

*Would you like to be able to see statistics of your usage of the application over time?*

*Were there any features you were expecting with the app that weren't there?*

*What kind of additional features would you expect from a premium version of the application?*

*How do you feel about the visual appearance of the app?*

*How do you find the flow and usability of the app?*

*What do you think about sharing user data with this app?*

*What kind of information would you be comfortable sharing with the app?*

### ***End***

*What do you think are the biggest advantages of the application?*

*What are the biggest disadvantages?*

*Is there anything you'd like to mention besides what we talked about today?*

## Appendix C. Adapted System Usability Scale

**1. I think that I would like to use this system frequently.**

<i>Strongly disagree</i>				<i>Neutral</i>				<i>Strongly agree</i>
1	2	3	4	5	6	7		

**2. I found the system unnecessarily complex.**

<i>Strongly disagree</i>				<i>Neutral</i>				<i>Strongly agree</i>
1	2	3	4	5	6	7		

**3. I thought the system was easy to use.**

<i>Strongly disagree</i>				<i>Neutral</i>				<i>Strongly agree</i>
1	2	3	4	5	6	7		

**4. I think that I would need the support of a technical person to be able to use this system.**

<i>Strongly disagree</i>				<i>Neutral</i>				<i>Strongly agree</i>
1	2	3	4	5	6	7		

**5. I found the various functions in this system were well integrated.**

<i>Strongly disagree</i>				<i>Neutral</i>				<i>Strongly agree</i>
1	2	3	4	5	6	7		

**6. I thought there was too much inconsistency in this system.**

<i>Strongly disagree</i>				<i>Neutral</i>				<i>Strongly agree</i>
1	2	3	4	5	6	7		

**7. I would imagine that most people would learn to use this system very quickly.**

<i>Strongly disagree</i>				<i>Neutral</i>				<i>Strongly agree</i>
1	2	3	4	5	6	7		

**8. I found the system very cumbersome to use.***Strongly disagree**Neutral**Strongly agree*

1                    2                    3                    4                    5                    6                    7

**9. I felt very confident using the system.***Strongly disagree**Neutral**Strongly agree*

1                    2                    3                    4                    5                    6                    7

**10. I needed to learn a lot of things before I could get going with this system.***Strongly disagree**Neutral**Strongly agree*

1                    2                    3                    4                    5                    6                    7