TU Delft

Delft University of Technology

A virtual sleepcoach for people suffering from insomnia

Horsch, Corine

DOI
10.4233/uuid:8da2aedf-ac26-4763-9271-66c6981007fe

Publication date
2016

Document Version
Final published version

Citation (APA)
Horsch, C. (2016). *A virtual sleepcoach for people suffering from insomnia*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:8da2aedf-ac26-4763-9271-66c6981007fe

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

A virtual sleepcoach for people suffering from insomnia

Corine Horsch

A virtual sleepcoach for people suffering from insomnia

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
23 november 2016 om 15:00 uur

door

Corine Hanelore Godefrida HORSCH
Master of Science in Human-Technolgy Interaction,
Eindhoven University of Technology, The Netherlands
Geboren te Eindhoven, The Netherlands

This dissertation has been approved by the
promotor: Prof. dr. M.A. Neerincx
copromoter: Dr. ir. W.P. Brinkman

Composition of the doctoral committee:

Rector Magnificus	chairperson
Prof. dr. M.A. Neerincx	Delft University of Technology, promotor
Dr. ir. W.P. Brinkman	Delft University of Technology, copromotor

Independent members:

Prof. dr. C.M. Jonker	Delft University of Technology
Prof. dr. H. de Ridder	Delft University of Technology
Prof. H. Oinas-Kukkonen	University of Oulu
Prof. dr. J.H.D.M. Westerink	Eindhoven University of Technology
Dr. ir. R.J. Beun	Utrecht University

Dr. J. Lancee of University of Amsterdam has contributed greatly to the preparation of this dissertation.

TU Delft

This research was funded by Philips and Technology Foundation STW, Nationaal Initiatief Hersenen en Cognitie NIHC under the Partnership program Healthy Lifestyle Solutions.

Printed by: Gildeprint Drukkerijen
Cover design by: Chantal Spaan
Copyright © 2016 by C.H.G. Horsch
ISBN 978-94-6186-732-2
An electronic version of this dissertation is available at http://repository.tudelft.nl

Table of Contents

Summary

People suffering from insomnia have problems falling asleep or staying asleep. Insomnia impairs people's daily life and their quality of life decreases. Approximately 10% of the population suffers from insomnia. The common treatment for insomnia is cognitive behavioural therapy for insomnia (CBT-I), mostly delivered by a therapist that people see once a week. A disadvantage of the current practice of insomnia treatment is the limited accessibility of insomnia treatment. Moreover, adherence to CBT-I exercises seems to be difficult. A virtual sleep coach that is provided through a smartphone might be a possible solution to both of these drawbacks. A virtual coach is never tired, never frustrated, never forgets things, and never gives up. Furthermore, it could improve accessibility, give tailored background information, offer personalized advice and feedback, monitor progress, provide support, and automatically track behaviour. Additionally, the majority of people in wealthy nations own a smartphone and emerging countries are expected to follow soon, making this type of intervention readily accessible to a large group of people. In short, a virtual sleep coach seems to be a good opportunity to improve traditional CBT-I. Concurrent to developing such a virtual sleep coach, answers to the question of how persuasive strategies can contribute to treatment adherence in an effective virtual sleep coach are explored.

The following concepts were investigated by several studies: adherence, persuasive strategies, and effectiveness. First, adherence rates were explored in a meta-analysis. This meta-analysis included 18 studies that researched technology-mediated sleep interventions. The data from those studies were retrieved and aggregated into an average that indicates that 52% of the participants adhered. The study confirmed the expected positive relationship between adherence and effectiveness. The meta-analysis was complemented with 15 semi-structured interviews about sleep support technologies and 6 focus groups regarding the to-be-developed virtual sleep coach. These qualitative studies were set-up to explore reasons for (non-)adherence and attitudes towards adherence-enhancing strategies. Participants expressed to rely on their own willpower for adherence, and that the concepts "users in control" and "doing it for your own sake" are important. However, there seem to be a discrepancy between the participant's perceived adherence rate and the adherence rate found in the meta-analysis. Because of this so called "adherence bias", the results from the interviews and focus groups should be interpreted with caution. A virtual sleep coach should be able to cope with this "adherence bias", and persuade users to accept adherence-enhancing strategies.

Two field experiments were conducted to investigate two different types of adherence-enhancing mechanisms. The first experiment investigated how to support users to get ready-to-change in a self-help setting. An interactive coach was developed to help users progress through the transtheoretical model (TTM) stages. This digital coach was compared to a paper workbook that contained the same exercises. The tool was a stage-matched self-reflection program that aimed to solve ambivalence. The experiment had a mixed design with within-subject pre-post measures, and between-subject paper

versus interactive tool conditions. Thirty-three participants were randomly assigned to one of the conditions, and were asked to work with their tool at least twice in a period of four weeks. Unfortunately, only about half of the participants met this requirement. Qualitative data revealed that users are ambivalent not only about their behaviour change, but also about the interventions and tools that support these change processes. The results suggest that non-adherence can still occur because of tool ambivalence, even though a virtual coach is stage-matched and uses persuasive strategies. In conclusion, a virtual coach should take the user's ambivalence towards itself into account, for example by reducing usage 'costs', increasing the perceived benefits, or tailoring the intervention to personal drives.

In the other field experiment two different types of reminders and their underlying principles were studied in relation to adherence. The first type of reminder was set by users themselves. This reminder was based on the idea that it increases the perceived self-empowerment of users, and thereby also increased adherence. The second type of reminder was automatically triggered based on events. This reminder reflected the idea that reminders sent at opportune moments will increase adherence. Both reminders and the underlying mechanisms self-empowerment and opportunity were studied in a within-subject experiment. Forty-five participants were randomly assigned to one of three conditions for one week each. They received no reminders, self-set reminders, or automatic reminders. The results showed that both reminders increase adherence. In addition, a mediation analysis showed that the effect of reminders on adherence could be partly explained by the perceived self-empowerment and opportuneness.

Lastly, the effectiveness of a virtual sleep coach that encompassed sleep restriction, a sleep diary, relaxation exercises, sleep hygiene, education, reminders and negotiation was studied. In a randomized controlled trial that included 151 participants and spanned 7 weeks the effects of the virtual sleep coach on various sleep measures was tested. The results show that the app had significant moderate effects on insomnia severity and sleep efficiency, which were the two main outcome measures. Furthermore, the majority of the other sleep measures also improved.

Future work could investigate other persuasive strategies, and study self-learning adaptive personalized virtual coaches. By adjusting persuasive strategies or the therapy to individuals, adherence and effectiveness of a virtual coach may be improved. Furthermore, adaptive personalisation can help to gain insights into the question of how persuasive strategies are effective. For example, the need for cognition might influence susceptibility for certain strategies. By studying adaptive personalized virtual coaches, these relationships might be disclosed.

In summary, adherence, persuasive strategies, and effectiveness were investigated in several empirical studies in order to answer the question in what way persuasive strategies can contribute to treatment adherence in an effective virtual sleep coach. The results show that an app can be effective in treating insomnia, and that adherence can, for example, be enhanced by reminders. However, adherence is still a topic of concern that should be studied more thoroughly.

Samenvatting

Mensen die lijden aan insomnie hebben problemen met in slaap vallen of doorslapen. Insomnie raakt mensen in hun dagelijks leven en hun kwaliteit van leven neemt af. Ongeveer 10% van de bevolking lijdt aan insomnie. De gebruikelijke behandeling van insomnie is cognitieve gedragstherapie voor insomnie (CGT-I) aangeboden door een therapeut die mensen één keer per week zien. Een nadeel van deze gebruikelijke praktijk is de beperkte toegankelijkheid van zo'n insomnie behandeling. Bovendien, lijkt therapietrouw aan CGT-I oefeningen lastig te zijn. Een virtuele slaap coach via een smartphone kan een mogelijke oplossing voor beide nadelen zijn. Een virtuele coach is nooit moe, nooit gefrustreerd, vergeet nooit de dingen, en geeft nooit op. Daarnaast zou de virtuele coach de toegankelijkheid tot insomnie behandeling kunnen verbeteren, achtergrondinformatie op maat kunnen geven, gepersonaliseerd advies en feedback kunnen geven, voortgang kunnen meten, ondersteuning kunnen bieden en automatisch gedrag kunnen bijhouden. Tevens bezit de meerderheid van de mensen in rijke landen een smartphone en groeilanden zullen naar verwachting spoedig volgen, waardoor dit type interventie gemakkelijk toegankelijk is voor een grote groep mensen. Kortom, een virtuele slaap coach lijkt een goede kans om de traditionele CGT-I te verbeteren. Tegelijkertijd met het ontwikkelen van een dergelijke virtuele slaap coach, worden antwoorden op de vraag hoe persuasieve strategieën kunnen bijdragen aan therapietrouw in een effectieve virtuele slaap coach onderzocht.

De volgende concepten zijn onderzocht in verschillende studies: therapietrouw, persuasieve strategieën en effectiviteit. Als eerste stap is therapietrouw onderzocht in een meta-analyse. Deze meta-analyse omvatte 18 studies die technologie-gemedieerde slaap interventies onderzochten. Uit de samengestelde data van deze studies bleek dat 52% van de deelnemers therapietrouw was. De meta-analyse bevestigt de verwachte positieve relatie tussen therapietrouw en effectiviteit. De studie werd aangevuld met 15 semigestructureerde interviews over slaap ondersteunende technologieën en 6 focusgroepen betrekking hebbend op de te ontwikkelen virtuele slaap coach. Deze kwalitatieve studies werden opgezet om de redenen achter slechte therapietrouw en attitudes ten opzichte van therapietrouw bevorderende strategieën te verkennen. Deelnemers brachten onder woorden dat ze vertrouwen op hun eigen wilskracht wat betreft therapietrouw. Daarnaast bleken de begrippen "de controle hebben" en " het voor je eigen bestwil doen" belangrijk te zijn. Er lijkt echter een discrepantie te bestaan tussen de deelnemers eigen waargenomen therapietrouw en de therapietrouw gevonden in de meta-analyse. Vanwege deze zogenoemde "therapietrouw bias", moeten de resultaten van de interviews en focusgroepen met voorzichtigheid worden geïnterpreteerd. Een virtuele slaap coach moet in staat zijn om met deze "therapietrouw bias" om te gaan en gebruikers overtuigen om therapietrouw bevorderende strategieën te accepteren.

Er zijn twee veldexperimenten uitgevoerd om twee verschillende types therapietrouw bevorderende mechanismen te onderzoeken. Het eerste experiment onderzocht hoe gebruikers ondersteund kunnen worden in een zelfhulp situatie om klaar-voor-verandering te zijn. Hiervoor is een interactieve coach ontwikkeld die gebruikers door de fasen van het transtheoretische model (TTM) helpt. Deze digitale coach werd vergeleken met een papieren werkboek die dezelfde oefeningen bevatte. De coach en het werkboek waren zelfreflectie tools gericht op het oplossen van ambivalentie en afgestemd op de TTM fase van de gebruiker. Het experiment had een gemengd ontwerp met per persoon pre-post metingen (within-subject), en verschillende condities tussen personen (between-subject): papier versus interactieve tool. Drieëndertig deelnemers werden willekeurig toegewezen aan een van de condities en gevraagd om tenminste tweemaal de tool te gebruiken in een periode van vier weken. Helaas bleek dat slechts ongeveer de helft van de deelnemers aan deze eis voldeed. Uit kwalitatieve data bleek dat gebruikers niet alleen ambivalent zijn over hun gedragsverandering, maar ook over de interventies en tools die deze verandering ondersteunen. De resultaten suggereren dat lage therapietrouw voor kan komen als gevolg van toolambivalentie, ook al is de virtuele coach afgestemd op de TTM-fase en maakt hij gebruik van persuasieve strategieën. Dientengevolge zou een virtuele coach de ambivalentie van de gebruiker aangaande de tool ook in aanmerking moeten nemen, bijvoorbeeld door het verminderen van de 'kosten' van het gebruik, het verhogen van de waargenomen voordelen, of het afstemmen van de interventie om persoonlijke drijfveren.

In het andere veldexperiment werden twee verschillende soorten herinneringen en de onderliggende principes onderzocht in relatie tot therapietrouw. Het eerste type herinnering werd door de gebruiker zelf ingesteld. De werking van dit type herinnering is gebaseerd op het idee dat de waargenomen *self-empowerment* van de gebruikers, en daarmee ook therapietrouw toeneemt. Het tweede type herinnering werd automatisch geactiveerd op basis van gebeurtenissen. Dit type herinnering weerspiegelde het idee dat herinneringen verzonden op geschikte momenten therapietrouw zullen doen toenemen. Zowel de herinneringen en de onderliggende mechanismen *self-empowerment* en geschiktheid werden onderzocht in een *within-subject* experiment. Vijfenveertig deelnemers werden willekeurig toegewezen aan één van drie één-week-durende condities. Ze kregen geen herinneringen, zelf ingestelde herinneringen of automatische herinneringen. De resultaten toonden aan dat beide herinneringen therapietrouw verhogen. Daarnaast toonde een mediatie analyse aan dat het effect van herinneringen op therapietrouw deels kan worden verklaard door de waargenomen *self-empowerment* en geschiktheid. Als laatste werd de effectiviteit van een virtuele slaap coach onderzocht. Deze slaap coach bestond uit slaaprestrictie, een slaap dagboek, ontspanningsoefeningen, slaaphygiëne, onderwijs, herinneringen en onderhandeling. In een *randomized controlled trial* werden de effecten van de virtuele slaap coach op 151 deelnemers gedurende 7 weken op diverse slaapmaten getest. De resultaten tonen aan dat de app middelgroot effect heeft op slaapefficiëntie en de mate van insomnie, welke de twee belangrijkste uitkomstmaten waren. Daarnaast zijn de meeste andere slaapmaten ook verbeterd.

Toekomstig onderzoek kan andere persuasieve strategieën en zelflerende adaptieve gepersonaliseerde virtuele coaches bestuderen. Door persuasieve strategieën of de therapie aan te passen aan de individu, kunnen therapietrouw en effectiviteit van een virtuele coach worden verbeterd. Daarnaast kan adaptieve personalisatie helpen om inzicht te verkrijgen in de vraag op welke manier persuasieve strategieën effectief zijn. Bijvoorbeeld, de behoefte aan kennis (need for cognition) kan de vatbaarheid voor bepaalde strategieën beïnvloeden. Door adaptieve gepersonaliseerde virtuele coaches te onderzoeken kunnen deze relaties onthuld worden.

Samengevat, therapietrouw, persuasieve strategieën en effectiviteit zijn onderzocht in verschillende empirische studies om de vraag op welke wijze persuasieve strategieën kunnen bijdragen aan therapietrouw in een effectieve slaap virtuele coach te beantwoorden. De resultaten tonen aan dat een app bij de behandeling van slapeloosheid effectief kan zijn en dat therapietrouw bijvoorbeeld kan worden verbeterd door herinneringen. Echter, therapietrouw is nog steeds een onderwerp van zorg dat grondiger moet worden bestudeerd.

1. Introduction

1.1. Insomnia and insomnia treatment

People who suffer from insomnia have difficulties with initiating or maintaining sleep (Morin, Barlow, & Dement, 1993). A review of the literature estimated that 9-15% of the western adult population is suffering from insomnia (Ohayon, 2002). This sleep disturbance significantly impairs people's daily functioning (American Psychiatric Association, 2013). Having insomnia may lead to personal suffering, such as feeling tired after a night's sleep, reduced quality of life, and vulnerability to depression (Baglioni et al., 2011; Rosekind & Gregory, 2010). In addition, insomnia leads to societal costs that might include reduced productivity and more sick leave from work (Daley, Morin, LeBlanc, Grégoire, & Savard, 2009; Rosekind & Gregory, 2010).

The common treatments for insomnia are pharmacotherapy and cognitive behavioural therapy for insomnia (CBT-I) (Dautovich, McNamara, Williams, Cross, & McCrae, 2010). Both treatments are effective (Riemann & Perlis, 2009). However, CBT-I is preferable, because CBT-I is equally effective in the short term and has more beneficial long-term effects than pharmacotherapy (Morin, Gaulier, Barry, & Kowatch, 1992; Perlis, Smith, Cacialli, Nowakowski, & Orff, 2003; Vincent & Lionberg, 2001). Generally, CBT-I consists of weekly sessions in which the focus lies on one or more of the following exercises: sleep restriction, stimulus control, relaxation, cognitive strategies, and sleep education and hygiene (Morin et al., 1993; Morin & Espie, 2003; Verbeek & Klip, 2005).

Sleep restriction (Spielman, Saskin, & Thorpy, 1987) tunes the time spent in bed to the actual sleep time. In practice, people fill in a sleep diary for approximately a week. A sleep diary contains the times people go to bed, fall asleep, wake up, get out of bed, and are awake during the night (Carney et al., 2012). Based on that data, an average total sleep time (TST) and an average total time in bed (TIB) is calculated, as well as the sleep efficiency (SE), which is the TST divided by the TIB. At the start of the sleep restriction exercise, people are only allowed to spend their TST in bed, until their SE reaches a certain threshold (usually 85%) (Kyle et al., 2015). When that threshold is reached, they are allowed some extra time in bed (usually 15-30 minutes) (Kyle et al., 2015). This process continues until SE cannot be improved any further. In general, the minimal bed time is less than the needed sleep time in the beginning of this exercise, because total time slept as reported in the diary is often underestimated by people suffering from insomnia (Bootzin & Epstein, 2011). Sleep restriction is strongly related to treatment outcome (Harvey, Inglis, & Espie, 2002; Miller et al., 2014; Morin, Culbert, & Schwartz, 1994). Unfortunately, sleep restriction is the treatment component that is least preferred by people suffering from insomnia (Vincent & Lionberg, 2001), and adhering to bedtime recommendations can be difficult (Riedel & Lichstein, 2001).

Stimulus control (Bootzin, 1972, 1979) aims at associating the bedroom with sleep (again). This implies that people are not allowed to do anything else in the bedroom except sleeping (and having sex). In practice, this means that people have to get out of bed when they lie awake for a while (originally 10 minutes)(Bootzin, Epstein, & Wood, 1991). They have to go to another room and are allowed to do a non-arousing

activity, such as reading a book or listening to music, until they feel really sleepy again; at that moment, they are allowed to go back to bed. Perceived barriers while doing sleep restriction are boredom, annoyance, and disturbing others (Vincent, Lewycky, & Finnegan, 2008). The less people perceive barriers to stimulus control, the better they adhere, and the better the outcomes are (Vincent et al., 2008). Studies have shown that stimulus control is one of the most effective single-component intervention for insomnia (e.g., Morin et al., 2006; Morin et al., 1999).

Relaxation exercises can help people to release muscle tension, and free up their minds. It is not always easy to accomplish such a relaxed state, but it could help to fall asleep quicker. Because hyperarousal has been seen as a determinant for insomnia, relaxation training has a long history in insomnia treatment (Bootzin & Epstein, 2011). Progressive muscle relaxation (Jacobson, 1938) is one of the most recommended relaxation exercises within sleep therapy (Bootzin & Epstein, 2011). It involves muscle tension release cycles, breathing control, and imagery. However, people could also benefit from other relaxation procedures, such as mindfulness, imagery exercises, and deep breathing, as long as people become relaxed (Morgenthaler et al., 2006).

The cognitive component (Belanger, Savard, & Morin, 2006; Harvey, 2002, 2005) in CBT-I aims at changing beliefs and attitudes regarding sleep. So, thoughts that are detrimental to sleep are tackled. For example, people might hold the dysfunctional belief 'I have to sleep for eight hours a night, otherwise I will not function well during the day'. Such a belief contributes to an anxiety about sleep and makes it difficult to fall asleep, often resulting in a vicious circle. In the cognitive exercises, people are made aware of their beliefs and their detrimental effect, and are challenged to test those beliefs (against the truth). By this experience, personal beliefs and attitudes can be changed into helpful thoughts regarding sleep, and thereby improve sleep (Edinger, Wohlgemuth, Radtke, Marsh, & Quillian, 2001; Morin, Blais, & Savard, 2002).

Sleep education and sleep hygiene (Hauri, 1977) increases people's knowledge about sleep and sleep habits. There is no consensus among sleep experts regarding sleep hygiene and education, and specific recommendation differ across studies (Stepanski & Wyatt, 2003). Sleep education encompasses the basic information about the nature of sleep, sleep needs, and consequences of sleep loss. This knowledge helps people to understand, accept, and comply to treatment recommendations. Sleep hygiene consists of recommendations regarding lifestyle for everyone to promote good sleep. Sleep hygiene consists of tips about food and drinks, the bed room, and behaviour that influences sleep. For example, caffeine, nicotine, alcohol, diet and exercises stimulate the central nervous system and thereby increase wakefulness. Sleep hygiene also encompasses knowledge about a good sleeping environment. Simply stated, people should have a comfortable bed, the temperature in the room should be good, air quality should be good, and the noise level should not be too high. A study that compared participants' liking of the CBT-I exercises before treatment, showed that sleep hygiene was liked better than the other CBT-I exercises (Vincent & Lionberg, 2001).

1.2. Treatment adherence

While CBT-I is the preferred treatment for insomnia, adherence to the CBT-I exercises can be difficult. For example, staying awake while being tired requires much effort. A quantitative review showed an adherence rate of 65.5% to sleep disorder treatments, the lowest adherence rate among 17 medical problems (like cancer, eye disorders, and cardio vascular diseases) studied in that review (DiMatteo, 2004). The 569 studies included in the review show an adherence rate of 75% across all treatments (DiMatteo, 2004). Moreover, the World Health Organization (WHO) recognizes the importance of adherence to health regimes, stating that, "Adherence is a primary determinant of the effectiveness of treatment" (World Health Organization, 2003). The impact of adherence on treatment outcome and the low adherence rate to sleep disorder treatments warrant further investigation into how adherence could be enhanced within an intervention in the context of insomnia therapy.

1.3. Persuasive strategies

Scholars have suggested various ideas to improve adherence to behaviour change interventions. One possible solution-path comprises persuasive strategies (e.g. Kelders, 2012). There are many persuasion strategies depending on the exhaustiveness, exclusivity, emphasis, and granularity of the 'search' (Kaptein, 2011). Various frameworks take different perspectives which can inspire designers and make them think about their own intervention and context in different ways. For example, Fogg (2003), Oinas-Kukkonen and Harjumaa (2008), and Consolvo, McDonald, and Landay (2009) take a technology perspective, whereas (Cialdini, 1993) takes more a marketing perspective.

Fogg (2003) defines persuasion as an attempt to change attitudes or behaviours or both (without using coercion or deception) (Fogg, 2003, p.15). Fogg can be seen as one of the 'founders' of the persuasive technology field and in his book he identified fifteen persuasion strategies suitable for computers (Fogg, 2003). His research is rather comprehensive and has been (partly) worked out in numerous research and development projects (e.g., Connelly et al., 2006). Fogg categorized persuasive strategies along 'the functional triad' which is based on how a technology is functioning (Fogg, 2003).The functional triad distinguishes three technology functions: technology as a tool, medium, or social actor. All have different underlying persuasive mechanisms. For example, tools may be persuasive because they can make target behaviour like filling in a sleep diary easier, whereas social actors might give positive feedback when a coachee gets out of bed while still sleepy. The strategies in which technology functions as a medium could persuade coachees by predicting or simulating how a treatment could work out for them in the future.

Oinas-Kukkonen and Harjumaa (2008) state that the strategies of Fogg are not directly usable for software implementations. Therefore, they proposed a systematic method to develop persuasive systems based on Fogg's strategies. Their Persuasive Systems Design (PSD) model contains three steps: designers should understand the key issues behind persuasive systems (step 1), designers should analyse the persuasion context (step 2), and designers should consider the system qualities (step 3). They comprehensively described the system qualities (step 3); the underlying principles of the system qualities are described, the belonging software requirements are written down, and examples of implementation are given. The PSD model divides the system qualities into four categories: task-, dialogue-, credibility-, and social support. The strategies within the first two categories (task and dialogue support) are based on the strategies of Fogg (2003). There are no fundamental differences; the descriptions deviate in the level of details. The underlying idea of credibility, the third category, is that a system that is more credible is more persuasive. Fogg (2003) discussed credibility as well, however not in his functional triad. For an insomnia intervention credibility might be assured by developing the intervention in narrow cooperation with sleep therapists to realize credible treatments and with coachees for perceived surface credibility. The last category is social support, which describes how to use social influence strategies to persuade the coachee. These social strategies could be used in an intervention by connecting the coachees with peers.

Consolvo, McDonald, and Landay (2009) derived guidelines for technologies that support behaviour change from implementations and the learned 'best practices'. The guidelines are 1) data abstraction and reflection, 2) unobtrusiveness, 3) public, 4) aesthetic, 5) positive, 6) controllable, 7) credibility, 8) historical, and 9) comprehensive. The method to base guidelines on prior examples is in line with the 8-step design approach of Fogg (2009), which states that designers should consider successful prior examples and imitate and expand on that success. Although these guidelines are not all about persuasion, they seem relevant 'learned lessons', since these guidelines aim at increasing the probability that technology will actually be used, and thereby adherence.

From a more marketing perspective six persuasive strategies were identified: reciprocity, liking, authority, consensus, scarcity, and consistency & commitment (Cialdini, 1993). Recently, these six strategies were used to emphasize the importance of tailoring in persuasion by means of persuasion profiles (Kaptein, 2011). A persuasion profile indicates how susceptible an individual is to a specific persuasion strategy. Based on the persuasion profile of a person a system can select and use the most effective persuasion strategy for that person. Advice in an intervention could for example be framed using authority or using consensus to persuade coachees to optimize their bedroom environment, depending on the persuasion profile of the coachee.

1.4. Virtual coach

A possible solution to support treatment adherence to CBT-I could be a virtual coach. In this thesis a virtual coach can be delivered in many formats, however, it should meet the following requirements: a) the virtual coach is a digital (programmed) coach, b) there is some sort of interaction possible with the virtual coach, and c) the virtual coach gives advice to the user. Advantages of a virtual coach are that it is never tired, never frustrated, never forgets things, and never gives up. For example, a virtual coach keeps sending reminders even if users do seldom comply to them. Potential other advantages of a virtual coach are improving accessibility of treatments, tailored background information, personalized advice and feedback, monitoring of progress, support via social networks, and automatic tracking (Klaassen, 2015). Accessibility can be improved because a virtual coach can be accessed anytime and anywhere. Thus, treatment can be supported across a variety of settings, broader than the typical 1 hour therapy session. A virtual coach also accommodate opportunities to tailor therapy. For instance, personal bedtimes can be calculated easily, preferences for relaxation exercises can be learned, or reminder texts can be personalised. Progress of users can be monitored precisely and continuously, even real-time if necessary.

The aforementioned benefits of virtual coaches could be effectuated through different media, e.g., via a computer, or smartphone. In 2015 there were more than 7 billion mobile phone subscriptions, while the world population was approximately 7.4 billion people (a penetration rate of 97%) (ITU, 2015). A global median of 43% of the mobile phone users own a smartphone (PRC, 2016). A strong correlation was found between the wealth of a nation and smartphone ownership. It is expected, however, that smartphone ownership will rapidly increase in emerging countries as well. Thus, delivering a virtual coach through a smartphone will expectedly give a vast majority of the world population access to treatment in the near future. Furthermore, according to studies conducted in Israel and Australia most smartphone owners are interested in monitoring their mental health via their mobile phones (Proudfoot et al., 2010; Torous, Friedman, & Keshavan, 2014). An enormous number of apps have been developed in recent years. Approximately, 6% of the total number of apps targeted mental health issues (Donker et al., 2013). Mental health apps target for example depression (Kauer et al., 2012; Ly, Carlbring, & Andersson, 2012; Watts et al., 2013), anxiety (Dagöö et al., 2014), and borderline personality disorder (Rizvi, Dimeff, Skutch, Carroll, & Linehan, 2011). A systematic review of mental health apps suggests that they have the potential to be effective (Donker et al., 2013). This conclusion, however, should be interpreted with caution, since the review included only 8 studies describing 5 apps. From the review it becomes clear that more research regarding mental health apps is needed. Nevertheless, mental health apps are promising, since, smartphones are widespread, people seem willing to accept mental health apps, they are potentially effective, and smartphones can utilize the advantages of a virtual coach. Therefore, researching a virtual coach that delivers CBT-I through a smartphone app is expected to be worthwhile.

1.5. Research question and hypotheses

A key challenge of a virtual coach is to provide therapy support in such a way that the coachees really adhere to the regimen of the personal therapy plan. The aim of this thesis is to explore in what ways adherence-enhancing mechanisms can contribute to treatment adherence in an effective virtual sleep coach. The following main question has driven the research presented in this thesis:

> *In what way can persuasive strategies contribute to improve treatment adherence to, and consequently the effectiveness of, a CBT-I-based virtual sleep coach?*

To answer this question the three concepts, adherence, persuasive strategies, and effectiveness that comprises this main question are investigated. The first concept is adherence. Adherence rates are important to be able to determine the effectiveness of a treatment. Only when users adhere to the treatment protocol, the outcome can be attributed to the intervention (Gould & Clum, 1993). For cognitive behavioural therapy various authors (e.g., Beun, 2012; Donkin et al., 2011) mention that treatment adherence is a problem. Moreover, reports about adherence to various internet-based interventions show mixed results. For example, Eysenbach (2005) gives a few examples of internet-based interventions with adherence rates ranging from 1% to 35%. So, the first thing to examine are the current adherence rates. Furthermore, it is important to explore why people do or do not adhere. In this thesis the position is taken that adherence and effectiveness in technology-mediated sleep interventions are related to each other, like they are in conventional treatments.

The second concept of the main question concerns persuasive strategies. Several ideas have been raised to improve adherence to behaviour change interventions (Beun, 2012; Fogg, 2003; Horsch, Brinkman, van Eijk, & Neerincx, 2012; Michie et al., 2013). One of the dominant theories in the health domain is the TransTheoretical Model (TTM). The TTM describes six stages a person can experience when changing behaviour (Prochaska & Velicer, 1997). Several studies (Noar, Benac, & Harris, 2007; Prochaska & Velicer, 1997) suggested that interventions need to be in accordance with people's readiness-to-change stage to be most effective. As most health interventions, CBT-I targets people who are ready to change, while most people are probably not ready to change yet. Research in other health domains estimated that approximately 80% of the people is not ready for change (Laforge, Velicer, Richmond, & Owen, 1999; Prochaska & Velicer, 1997). Not being ready for change might be related to non-adherence. One strategy that can support people with progressing through the stages of change is Motivational Interviewing (MI) (Resnicow et al., 2002). In several studies computerized MI interventions were developed to motivate participants to be more physical active (Bickmore, Schulman, & Sidner, 2013; Blanson Henkemans et al., 2009; Di Noia, Contento, & Prochaska, 2008; Friederichs, Bolman, Oenema, Verboon, & Lechner, 2016). Moreover, one study concerning fruit and vegetable intake showed

that participants can move to a higher TTM stage when supported by computerized MI intervention (Di Noia et al., 2008). In this thesis the position is taken that a self-reflection tool can help people progress through the readiness-to-change stages about their sleep problem.

Another reason for none-adherence is forgetfulness (Donkin & Glozier, 2012; Horsch, Lancee, Beun, Neerincx, & Brinkman, 2015). Sending reminders is a simple method that could decrease forgetfulness (Donkin & Glozier, 2012; Krishna, Boren, & Balas, 2009). Earlier research showed that mobile text reminders increase show-up rates for gastrointestinal endoscopy (Deng et al., 2015), breast cancer screening (Vidal et al., 2014), sunscreen use (Armstrong et al., 2009), and logging food intake (Bentley & Tollmar, 2013). In this thesis the position is taken that reminders sent by a virtual coach can also be effective in sleep interventions. These reminders can be implemented in various ways, and thereby function because of different underlying psychological principles. For example, users could set reminders themselves, or reminders could be sent automatically. Self-set reminders could work, because they increase the feeling of self-empowerment, and users might know best when they have time to perform an activity. Automatic reminders could be implemented in various ways. For instance, they could be based on the Capability-Opportunity-Motivation-Behaviour model (Michie, van Stralen, & West, 2011). This model suggests that if people are capable and motivated to exhibit a behaviour, a reminder at an opportune moment improves the change a person will exhibit this behaviour. This thesis explores if these underlying principles could explain their effectiveness.

The third concept concerns the effectiveness of a virtual coach. It is becoming more common to offer self-help interventions via the internet. Computerized treatments have a few advantages compared to face-to-face treatments. Computerized treatments can potentially save costs, because less time is needed from therapists. Additionally, the treatment can be offered to a larger number of people (Thorndike et al., 2008). CBT-I delivered via a smartphone has the same advantages as computerized CBT-I, but could theoretically surpass those advantages. Smartphones are ubiquitous, unobtrusive, and intimate (Klasnja & Pratt, 2012). Furthermore, smartphones are rich of sensors, computationally powerful, and remotely accessible. These properties provide opportunities for personalisation, ecological momentary access, and real-time tracking (Kaltenthaler & Cavanagh, 2010; Konrath & Yan, 2015). Moreover, a recent meta-analysis (Zachariae, Lyby, Ritterband, & O'Toole, 2016) demonstrated that internet-delivered CBT-I showed large treatment effects (Cohen's d = 1.0) on the Insomnia Severity Index. The position is taken that insomnia treatment can be effectively delivered by an fully automated smartphone app.

To conclude this section, from the main research question and the concepts introduced above, it is now possible to derive four hypotheses that are tested in this thesis:

H1: There is a positive relationship between adherence to a technology-mediated sleep intervention and the effectiveness of that intervention.

H2: A self-reflection tool can help people progress through the readiness-to-change stages.

H3: Computer-generated reminders increase adherence rates in technology-mediated sleep interventions.

H4: A fully automated virtual sleep coach app, encompassing sleep restriction, sleep diary, relaxation exercises, sleep hygiene, education, reminders and negotiation, is clinically effective in reducing insomnia.

1.6. Research approach

The first hypothesis about the relationship between effectiveness and adherence to technology-mediated sleep products was studied in three ways. First, a meta-analysis explored adherence rates of technology-mediated insomnia therapy across different studies reported in the literature. Several databases were queried, and the data of 18 studies were retrieved and aggregated to find an average adherence rate. In addition, the relationship between adherence rates and effectiveness was explored and confirmed. Another important viewpoint is that of the users. Eventually, users have to use and adhere to the virtual sleep coach, so this can make it, or break it. Therefore, 15 semi-structured interviews about sleep support technologies were conducted to investigate perceived adherence. The interviews included both adherence-related questions and questions regarding the factors of the Unified Theory of Acceptance and Use of Technology model. The transcriptions of the interviews were categorized to identify common themes and discover trends. Thirdly, 12 scenarios and 72 claims were written about the usage of a virtual sleep coach with explicitly scripted adherence-enhancing mechanisms. Well-known behaviour change theories and persuasive strategies were used to guide the ideas for improving adherence. The scenarios and claims were discussed in six different focus groups consisting of potential users (n=15), sleep experts (n=7), and coaches (n=9). The sessions were videotaped, transcribed, and summarized. The summaries were again categorized to identify common themes and discover trends. More details can be found in chapter two.

The second hypothesis about supporting users to get ready-to-change in a self-help setting was studied with a field experiment. The experiment was set up to find out if people can progress from one to another TTM-stage, and if an interactive coach was better able to help people in that process than a paper workbook. The experiment had a mixed design with within-subject pre-post measures, and between-subject a paper

versus an interactive tool. For this experiment a stage-matched self-help tool concerning sleep problems that helped participants to structure and reflect on their thoughts was developed. The tool was based on motivational interviewing principles, and consisted of seven chapters and 34 exercises. The 33 participants were randomly assigned to the paper or interactive tool, which were sent to their homes. The requirement was to work at least twice with the tool within a period of four weeks. More details and the results of this study can be found in chapter three.

The third hypothesis was explored with another field experiment. Two different types of reminders and the underlying principles were studied in relation to adherence. One type of reminder was set by users themselves. This user-based reminder reflected the idea that increasing self-empowerment increases adherence. The second type of reminder was event-based, and reflected the idea that reminding people at opportune moments increases adherence. The experiment had a within-subjects design with 45 participants who were exposed to three conditions during a total time of three weeks. In one condition participants received no reminders to perform targeted behaviour, in the other condition participants set the reminders themselves, and the last condition consisted of automatic event-based reminders. The order of the three conditions was counter-balanced across the participants. More details and the results can be found in chapter four.

The fourth hypothesis regarding the effectiveness of a sleep app was tested in a randomized controlled trial. This study had a between subject design with two arms: a waiting-list condition and a intervention condition with pre-, post-, and 3 month follow-up measures. 151 participants met the inclusion criteria and were randomly assigned to the app (n = 74), or a waiting-list condition (n = 79). The app consisted of a sleep diary, a relaxation exercise, sleep restriction, and sleep hygiene and education. It was fully automated and spanned maximal 7 weeks. The main measurements were the Insomnia Severity Index and sleep efficiency. The latter was measured with a separate 7-day online diary. More details and the results can be found in chapter five.

Lastly, chapter six draws conclusions on the way in which persuasive mechanisms contribute to treatment adherence in an effective virtual sleep coach. It also reflects on existing behaviour change theories and persuasive strategies. Furthermore, some recommendations regarding the development of a virtual sleep coach based on the most important lessons learned are provided.

1.7. References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed ed.). Arlington, VA: American Psychiatric Publishing.

Armstrong, A. W., Watson, A. J., Makredes, M., Frangos, J. E., Kimball, A. B., & Kvedar, J. C. (2009). Text-message reminders to improve sunscreen use: a randomized, controlled trial using electronic monitoring. Archives of dermatology, 145(11), 1230-1236.

Baglioni, C., Battagliese, G., Feige, B., Spiegelhalder, K., Nissen, C., Voderholzer, U., . . . Riemann, D. (2011). Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. J Affect Disord, 135(1), 10-19.

Belanger, L., Savard, J., & Morin, C. M. (2006). Clinical management of insomnia using cognitive therapy. Behavioral Sleep Medicine, 4(3), 179-202.

Bentley, F., & Tollmar, K. (2013). The Power of Mobile Notifications to Increase Wellbeing Logging Behavior. Paper presented at the ACM SIGCHI International Conference on Human Factors in Computing Systems.

Beun, R. J. (2012). Persuasive strategies in mobile insomnia therapy: alignment, adaptation, and motivational support. Personal and Ubiquitous Computing, 1-9. doi: 10.1007/s00779-012-0586-2

Bickmore, T. W., Schulman, D., & Sidner, C. (2013). Automated interventions for multiple health behaviors using conversational agents. Patient Educ Couns, 92(2), 142-148.

Blanson Henkemans, O. A., van der Boog, P. J. M., Lindenberg, J., van der Mast, C. A. P. G., Neerincx, M. A., & Zwetsloot-Schonk, B. J. H. M. (2009). An online lifestyle diary with a persuasive computer assistant providing feedback on self-management. Technology and Health Care, 17(3), 253-267.

Bootzin, R. R. (1972). Stimulus control treatment for insomnia. Paper presented at the APA 80th Annual Convention, Honolulu, HI, September 2-8, 1972.

Bootzin, R. R. (1979). Effects of self-control procedures for insomnia. American Journal of Clinical Biofeedback, 2(2), 70-77.

Bootzin, R. R., Epstein, D., & Wood, J. M. (1991). Stimulus control instructions Case studies in insomnia (pp. 19-28): Springer.

Bootzin, R. R., & Epstein, D. R. (2011). Understanding and treating insomnia. Annual Review of Clinical Psychology, 7, 435-458.

Carney, C. E., Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Krystal, A. D., Lichstein, K. L., & Morin, C. M. (2012). The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep, 35(2), 287.

Cialdini, R. (1993) Influence, Science and Practice. HarperCollings College, New York.

Connelly, K., Faber, A., Rogers, Y., Siek, K., Toscos, T. (2006) Mobile applications that empower people to monitor their personal health. Elektrotechnik und informationstechnik, 123, 124-128.

Consolvo, S., McDonalde, D., Landay, J. (2009) Theory-Driven Design Strategies for Technologies that Support Behavior Change in Everyday Life. CHI 2009 - Creative Thought and Self-Improvement, April 6th, 2009, 405-414.

Dagöö, J., Asplund, R. P., Bsenko, H. A., Hjerling, S., Holmberg, A., Westh, S., . . . Furmark, T. (2014). Cognitive behavior therapy versus interpersonal psychotherapy for social anxiety disorder delivered via smartphone and computer: A randomized controlled trial. Journal of anxiety disorders, 28(4), 410-417.

Daley, M., Morin, C. M., LeBlanc, M., Grégoire, J.-P., & Savard, J. (2009). The economic burden of insomnia: direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms, and good sleepers. Sleep, 32(1), 55.

Dautovich, N. D., McNamara, J., Williams, J. M., Cross, N. J., & McCrae, C. S. (2010). Tackling sleeplessness: Psychological treatment options for insomnia. Nature and Science of Sleep, 2, 23.

Deng, X., Wang, Y., Zhu, T., Zhang, W., Yin, Y., & Ye, L. (2015). Short Message Service (SMS) can Enhance Compliance and Reduce Cancellations in a Sedation Gastrointestinal Endoscopy Center: A Prospective Randomized Controlled Trial. Journal of medical systems, 39(1), 1-11.

Di Noia, J., Contento, I. R., & Prochaska, J. O. (2008). Computer-mediated intervention tailored on transtheoretical model stages and processes of change increases fruit and vegetable consumption among urban African-American adolescents. American Journal of Health Promotion, 22(5), 336-341.

DiMatteo, M. R. (2004). Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. Medical care, 42(3), 200-209.

Donker, T., Petrie, K., Proudfoot, J., Clarke, J., Birch, M.-R., & Christensen, H. (2013). Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res, 15(11).

Donkin, L., Christensen, H., Naismith, S. L., Neal, B., Hickie, I. B., & Glozier, N. (2011). A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res, 13(3).

Donkin, L., & Glozier, N. (2012). Motivators and Motivations to Persist With Online Psychological Interventions: A Qualitative Study of Treatment Completers. J Med Internet Res, 14(3).

Edinger, J. D., Wohlgemuth, W. K., Radtke, R. A., Marsh, G. R., & Quillian, R. E. (2001). Does cognitive-behavioral insomnia therapy alter dysfunctional beliefs about sleep? Sleep: Journal of Sleep and Sleep Disorders Research.

Eysenbach, G. (2005). The law of attrition. J Med Internet Res, 7(1).

Fogg, B. J. (2003). Persuasive technology: using computers to change what we think and do. San Francisco: Morgan Kaufmann Publishers.

Friederichs, S., Bolman, C., Oenema, A., Verboon, P., & Lechner, L. (2016). Exploring the working mechanisms of a web-based physical activity intervention, based on self-determination theory and motivational interviewing. Internet Interventions, 3, 8-17.

Gould, R. A., & Clum, G. A. (1993). A meta-analysis of self-help treatment approaches. Clinical Psychology Review, 13(2), 169-186.

Harvey, A. G. (2002). A cognitive model of insomnia. Behaviour research and therapy, 40(8), 869-893.

Harvey, A. G. (2005). A cognitive theory and therapy for chronic insomnia. Journal of Cognitive Psychotherapy, 19(1), 41-59.

Harvey, L., Inglis, S. J., & Espie, C. A. (2002). Insomniacs' reported use of CBT components and relationship to long-term clinical outcome. Behaviour research and therapy, 40(1), 75-83.

Hauri, P. (1977). Current concepts: the sleep disorders. Kalamazoo, MI: The Upjohn Company.

Horsch, C., Brinkman, W. P., van Eijk, R., & Neerincx, M. (2012). Towards the usage of persuasive strategies in a virtual sleep coach. Paper presented at the UK HCI 2012 Workshop on People, Computers & Psychotherapy, Birmingham.

Horsch, C. H. G., Lancee, J., Beun, R. J., Neerincx, M. A., & Brinkman, W. P. (2015). Adherence to Technology-mediated Insomnia Treatment: A meta-analysis, interviews with users, and focus groups with users and experts. Journal of Medical Internet Research, 17(9), e214. doi: 10.2196/jmir.4115

ITU. (2015). ICT facts and figures. Geneva, Switzerland: International Telecommunication Union.

Jacobson, E. (1938). Progressive relaxation.

Kaltenthaler, E., & Cavanagh, K. (2010). Computerised cognitive behavioural therapy and its uses. Progress in Neurology and Psychiatry, 14(3), 22-29.

Kaptein, M. (2011) Personalized Persuasion in Ambient Intelligence. PhD Thesis, Eindhoven University of Technology, The Netherlands.

Kaptein, M., Markopoulos, P., de Ruyter, B., & Aarts, E. (2015). Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles. International Journal of Human-Computer Studies, 77, 38-51.

Kauer, S. D., Reid, S. C., Crooke, A. H. D., Khor, A., Hearps, S. J. C., Jorm, A. F., . . . Patton, G. (2012). Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. J Med Internet Res, 14(3), e67.

Kelders, S.M. (2012). understanding adherence to web-based interventions. PhD Thesis, University of Twente, The Netherlands.

Klaassen, R. (2015). HCI perspectives on behavior change support systems: University of Twente.

Klasnja, P., & Pratt, W. (2012). Healthcare in the pocket: mapping the space of mobile-phone health interventions. Journal of Biomedical Informatics, 45(1), 184-198.

Konrath, S., & Yan, Z. (2015). Positive technology. Using mobile phones for psychosocial interventions. Encyclopedia of Mobile Phone Behavior.

Krishna, S., Boren, S. A., & Balas, E. A. (2009). Healthcare via cell phones: a systematic review. Telemedicine and e-Health, 15(3), 231-240.

Kyle, S. D., Aquino, M. R. J., Miller, C. B., Henry, A. L., Crawford, M. R., Espie, C. A., & Spielman, A. J. (2015). Towards standardisation and improved understanding of sleep restriction therapy for insomnia disorder: A systematic examination of CBT-I trial content. Sleep Medicine Reviews, 23, 83-88.

Laforge, R. G., Velicer, W. F., Richmond, R. L., & Owen, N. (1999). Stage distributions for five health behaviors in the United States and Australia. Preventive medicine, 28(1), 61-74.

Ly, K. H., Carlbring, P., & Andersson, G. (2012). Behavioral activation-based guided self-help treatment administered through a smartphone application: study protocol for a randomized controlled trial. Trials, 13(1), 62.

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., . . . Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. Annals of behavioral medicine, 46(1), 81-95.

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implementation Science, 6(1), 42.

Miller, C. B., Espie, C. A., Epstein, D. R., Friedman, L., Morin, C. M., Pigeon, W. R., . . . Kyle, S. D. (2014). The evidence base of sleep restriction therapy for treating insomnia disorder. Sleep Medicine Reviews, 18(5), 415-424.

Morgenthaler, T., Kramer, M., Alessi, C., Friedman, L., Boehlecke, B., Brown, T., . . . Owens, J. (2006). Practice parameters for the psychological and behavioral treatment of insomnia: an update. An American Academy of Sleep Medicine report. SLEEP-NEW YORK THEN WESTCHESTER-, 29(11), 1415-1419.

Morin, C., Blais, F., & Savard, J. (2002). Are changes in beliefs and attitudes about sleep related to sleep improvements in the treatment of insomnia? Behaviour research and therapy, 40(7), 741-752.

Morin, C. M., Barlow, D. H., & Dement, W. C. (1993). Insomnia: Psychological assessment and management: Guilford Press New York.

Morin, C. M., Bootzin, R. R., Buysse, D. J., Edinger, J. D., Espie, C. A., & Lichstein, K. L. (2006). Psychological and behavioral treatment of insomnia: update of the recent evidence (1998-2004). SLEEP-NEW YORK THEN WESTCHESTER-, 29(11), 1398.

Morin, C. M., Culbert, J. P., & Schwartz, S. M. (1994). Nonpharmacological interventions for insomnia: a meta-analysis of treatment efficacy. The American Journal of Psychiatry; The American Journal of Psychiatry.

Morin, C. M., & Espie, C. A. (2003). Insomnia: A clinician's guide to assessment and treatment (Vol. 1): Springer.

Morin, C. M., Gaulier, B., Barry, T., & Kowatch, R. A. (1992). Patients' acceptance of psychological and pharmacological therapies for insomnia. Sleep: Journal of Sleep Research & Sleep Medicine.

Morin, C. M., Hauri, P. J., Espie, C. A., Spielman, A. J., Buysse, D. J., & Bootzin, R. R. (1999). Nonpharmacologic treatment of chronic insomnia. An American Academy of Sleep Medicine review. Sleep, 22(8), 1134-1156.

Noar, S. M., Benac, C. N., & Harris, M. S. (2007). Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull, 133(4), 673-693. doi: 10.1037/0033-2909.133.4.673

Ohayon, M. M. (2002). Epidemiology of insomnia: what we know and what we still need to learn. Sleep medicine reviews, 6(2), 97-111.

Oinas-Kukkonen, H., Harjumaa, M. (2008) A Systematic Framework for Designing and Evaluating Persuasive Systems. Persuasive 2008, Oulu, Finland, June 4-6. Springer-Verlag, Berlin.

Perlis, M. L., Smith, M. T., Cacialli, D., Nowakowski, S., & Orff, H. (2003). On the comparability of pharmacotherapy and behavior therapy for chronic insomnia: commentary and implications. Journal of psychosomatic research, 54(1), 51-59.

PRC. (2016). Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies. : Pew Research Center.

Prochaska, J. O., & Velicer, W. F. (1997). The transtheoretical model of health behavior change. American journal of health promotion, 12(1), 38-48.

Proudfoot, J. G., Parker, G. B., Pavlovic, D. H., Manicavasagar, V., Adler, E., & Whitton, A. E. (2010). Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. J Med Internet Res, 12(5), e64.

Resnicow, K., DiIorio, C., Soet, J. E., Borrelli, B., Hecht, J., & Ernst, D. (2002). Motivational interviewing in health promotion: It sounds like something is changing. Health Psychology, 21(5), 444-451.

Riedel, B. W., & Lichstein, K. L. (2001). Strategies for evaluating adherence to sleep restriction treatment for insomnia. Behaviour research and therapy, 39(2), 201-212.

Riemann, D., & Perlis, M. L. (2009). The treatments of chronic insomnia: a review of benzodiazepine receptor agonists and psychological and behavioral therapies. Sleep Medicine Reviews, 13(3), 205-214.

Rizvi, S. L., Dimeff, L. A., Skutch, J., Carroll, D., & Linehan, M. M. (2011). A pilot study of the DBT coach: an interactive mobile phone application for individuals with borderline personality disorder and substance use disorder. Behavior therapy, 42(4), 589-600.

Rosekind, M. R., & Gregory, K. B. (2010). Insomnia risks and costs: health, safety, and quality of life. Am J Manag Care, 16(8), 617-626.

Spielman, A. J., Saskin, P., & Thorpy, M. J. (1987). Treatment of chronic insomnia by restriction of time in bed. Sleep: Journal of Sleep Research & Sleep Medicine.

Stepanski, E. J., & Wyatt, J. K. (2003). Use of sleep hygiene in the treatment of insomnia. Sleep Medicine Reviews, 7(3), 215-225.

Thorndike, F. P., Saylor, D. K., Bailey, E. T., Gonder-Frederick, L., Morin, C. M., & Ritterband, L. M. (2008). Development and perceived utility and impact of an internet intervention for insomnia. E-journal of applied psychology: clinical and social issues, 4(2), 32.

Torous, J., Friedman, R., & Keshavan, M. (2014). Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR mHealth and uHealth, 2(1), e2.

Verbeek, I., & Klip, E. C. (2005). Slapeloosheid [Insomnia]. Amsterdam: Boom.

Vidal, C., Garcia, M., Benito, L., Milà, N., Binefa, G., & Moreno, V. (2014). Use of text-message reminders to improve participation in a population-based breast cancer screening program. Journal of medical systems, 38(9), 1-7.

Vincent, N., Lewycky, S., & Finnegan, H. (2008). Barriers to engagement in sleep restriction and stimulus control in chronic insomnia. Journal of Consulting and Clinical Psychology, 76(5), 820.

Vincent, N., & Lionberg, C. (2001). Treatment preference and patient satisfaction in chronic insomnia. Sleep.

Watts, S., Mackenzie, A., Thomas, C., Griskaitis, A., Mewton, L., Williams, A., & Andrews, G. (2013). CBT for depression: a pilot RCT comparing mobile phone vs. computer. BMC Psychiatry, 13(1), 1.

World Health Organization, W. (2003). Adherence to long-term therapies: evidence for action. Geneva, Switzerland: World Health Organization.

Zachariae, R., Lyby, M. S., Ritterband, L. M., & O'Toole, M. S. (2016). Efficacy of internet-delivered cognitive-behavioral therapy for insomnia–A systematic review and meta-analysis of randomized controlled trials. Sleep Medicine Reviews, 30, 1-10.

2. Adherence to technology-mediated insomnia treatment: A meta-analysis, interviews, and focus groups

Abstract

Background - Several technologies have been proposed to support the reduction of insomnia complaints. A user-centered assessment of these technologies could provide insight into underlying factors related to treatment adherence.

Objective - Gaining insight into adherence to technology-mediated insomnia treatment as a solid base for improving those adherence rates by applying adherence-enhancing strategies.

Methods - Adherence to technology-mediated sleep products was studied in three ways. First, a meta-analysis was performed to investigate adherence rates in technology-mediated insomnia therapy. Several databases were queried for technology-mediated insomnia treatments. After inclusion and exclusion steps, data from 18 studies were retrieved and aggregated to find an average adherence rate. Next, 15 semistructured interviews about sleep-support technologies were conducted to investigate perceived adherence. Lastly, several scenarios were written about the usage of a virtual sleep coach that could support adherence rates. The scenarios were discussed in six different focus groups consisting of potential users (n = 15), sleep experts (n = 7), and coaches (n = 9).

Results - From the meta-analysis, average treatment adherence appeared to be approximately 52% (95% *CI*: 43%-61%) for technology-mediated insomnia treatments. This means that, on average, half of the treatment exercises were not executed, suggesting there is a substantial need for adherence and room for improvement in this area. However, the users in the interviews believed they adhered quite well to their sleep products. Users mentioned relying on personal commitment (i.e., willpower) for therapy adherence. Participants of the focus groups reconfirmed their belief in the effectiveness of personal commitment, which they regarded as more effective than adherence-enhancing strategies.

Conclusions - Although adherence rates for insomnia interventions indicate extensive room for improvement, users might not consider adherence to be a problem; they believe willpower to be an effective adherence strategy. A virtual coach should be able to cope with this "adherence bias" and persuade users to accept adherence-enhancing strategies, such as reminders, compliments, and community building.

Chapter published as:

Horsch, C. H. G., Lancee, J., Beun, R. J., Neerincx, M. A., & Brinkman, W. P. (2015). Adherence to Technology-mediated Insomnia Treatment: A meta-analysis, interviews with users, and focus groups with users and experts. *Journal of Medical Internet Research, 17*(9), e214.

2.1. Introduction

People who suffer from insomnia have difficulties with initiating sleep, maintaining sleep, or early-morning awakenings, and this sleep disturbance significantly impairs their daily functioning (American Psychiatric Association, 2013). Having insomnia may lead to personal suffering, such as feeling tired after a night's sleep, reduced quality of life, and vulnerability to depression (Baglioni et al., 2011; Rosekind & Gregory, 2010). In addition, insomnia leads to societal costs that might include reduced productivity and more sick leave from work (Daley, Morin, LeBlanc, Grégoire, & Savard, 2009; Ohayon, 2002; Rosekind & Gregory, 2010). A review of the literature showed that about 9% to 15% of the western adult population suffers from insomnia symptoms and the daytime consequences thereof (Ohayon, 2002).

Although the consequences of insomnia may be severe and prevalence is substantial, only a few people seek treatment (Ancoli-Israel & Roth, 1999; Benca, 2005; Morin, LeBlanc, Daley, Gregoire, & Merette, 2006). When help is sought, insomnia is most commonly treated with pharmacotherapy (Benca, 2005). However, cognitive behavioral therapy for insomnia (CBT-I) is preferable, because CBT-I is equally effective in the short term and has more beneficial long-term effects than pharmacotherapy (Morin, Gaulier, Barry, & Kowatch, 1992; Perlis, Smith, Cacialli, Nowakowski, & Orff, 2003; Vincent & Lionberg, 2001). Generally, CBT-I consists of weekly sessions in which the focus lies on one or more of the following exercises: sleep restriction, stimulus control, relaxation, cognitive strategies, and sleep hygiene (Morin & Espie, 2003).

Although CBT-I is effective, there is a lack of knowledge and accessibility regarding this type of therapy (Morin, 1993). General practitioners are often not aware of the existence of CBT-I, and neither is the general public (Morin, 1993). In addition, there are too few sleep therapists to help all people with insomnia (Morin, Beaulieu-Bonneau, LeBlanc, & Savard, 2005). In order to increase the availability and accessibility of CBT-I, Espie and colleagues (2012) suggested a stepped model with Internet-based treatment as a first option. A meta-analysis about computerized CBT-I (CCBT-I) concluded that this therapy is a moderately effective self-help intervention for insomnia (Cheng & Dizon, 2012). Nonetheless, adherence to insomnia and other technology-mediated treatments is often mentioned as a serious problem (Beun, 2012; Donkin et al., 2011; Eysenbach, 2005).

The World Health Organization (WHO) recognizes the importance of adherence to health regimes in general. They stated, "Adherence is a primary determinant of the effectiveness of treatment" (World Health Organization, 2003). In agreement with the WHO statement, Gould and Clum (1993) found—in their meta-analysis of self-help treatments—that better adherence to a treatment improves the treatment effectiveness. They found that the effect size was three times higher for studies that

had 75% to 100% adherence than for studies with adherence rates lower than 75%. The impact of adherence on treatment outcomes therefore warrants further investigation into how we could enhance adherence within an intervention in the context of insomnia therapy.

Various authors, for example, Beun (2012) and Donkin and colleagues (2011), mention that treatment adherence is a problem for CBT in general. Reports about adherence to various Internet-based interventions show mixed results. For example, Eysenbach (2005) gives a few examples in his "law of attrition" of Internet-based interventions with adherence rates ranging from 1% to 35%. Interestingly, a meta-analysis about CCBT-I reported an average adherence rate of 78% for the six studies they included (Cheng & Dizon, 2012). However, they did not make a distinction between treatment adherence and experimental compliance, that is, the proportion of the experimental assessments, such as questionnaires, that are completed. Thus, decisive conclusions on the exact adherence rates cannot be made.

The studies in this paper are conducted in the context of the Sleepcare project (Beun, Griffioen-Both, Ahn, Fitrianie, & Lancee, 2014; Horsch, Brinkman, van Eijk, & Neerincx, 2012), which aims at the development of a virtual sleep coach that delivers personalized, automated sleep therapy via a mobile phone. A key challenge of this e-coach is to provide therapy support in such a way that the coachees really adhere to the regimen of the personal therapy plan. In this paper, we use the generic term coachee—instead of client, patient, user, etc—to refer to both patients and nonpatients who seek help to address their health issues. The first step in the development of a virtual sleep coach that meets this adherence challenge is the analysis of current adherence rates, current adherence-enhancing strategies, and coachees' willingness to accept those strategies. Therefore, we conducted a meta-analysis about adherence rates in technology-mediated sleep interventions; interviewed coachees about their adherence to existing sleep-supporting technology; and discussed adherence-enhancing strategies in a to-be-developed virtual sleep coach among focus groups with potential users, sleep experts, and coaches. This complementary analysis approach provided new insights on how a virtual coach can support coachees to adhere to sleep therapy (i.e., the needs and constraints).

2.1.1. Study I: Meta-Analysis Adherence Rates

In order to determine whether a certain outcome is related to a treatment, adherence rates must be measured. Otherwise, it cannot be claimed that the outcome was caused by the intervention (Gould & Clum, 1993). Capturing adherence data is relatively easy in technology-mediated interventions (Donkin et al., 2011). However, as there is currently no standard adherence measure (Donkin et al., 2011; Edinger & Means, 2005; Vermeire, Hearnshaw, Van Royen, & Denekens, 2001), various measures are used. A review (Donkin et al., 2011) of adherence in e-therapies found the following adherence measures: number of log-ins, completed modules, number of visits/posts to a forum, pages viewed/printed, and self-reported measures. Other measures that have

been suggested are the usage time of the technology (Christensen, Griffiths, & Farrer, 2009) and reports by a spouse or related others (Edinger & Means, 2005). Different measures have different advantages and disadvantages. For example, time spent using the technology is an objective measure. However, time spent is presumably influenced by cognitive ability, reading speed, familiarity with the technology, etc (Donkin et al., 2011). Therefore, time spent does not necessarily represent treatment adherence. Moreover, there is a difference in passively using material (i.e., reading, listening, watching) and actively applying this material (i.e., performing the exercises) (Gould & Clum, 1993).

First, it is important to distinguish between at least two concepts: *treatment adherence* and *experiment compliance*. Treatment adherence refers to the extent a coachee processes and applies the content of the treatment (as provided by the coach), whereas experiment compliance refers to the coachees' completion of the experimental assessments. Other researchers have also made this distinction. For example, Christensen and colleagues (Christensen et al., 2009) respectively use the terms adherence (experience content) and dropout (research trial protocol), whereas Hebert and colleagues (Hebert, Vincent, Lewycky, & Walsh, 2010) respectively call it nonusage attrition and study attrition. Treatment adherence and experiment compliance might be related, but to our knowledge no information about this relationship has been reported in the literature.

2.1.2. Study II: User Adherence to Existing Sleep-Supporting Technology

After analyzing reported adherence rates to technology-mediated sleep treatment in the literature, the next step was to study coachees' reasons why they do or do not adhere to technology-mediated sleep interventions. To do so, interviews were conducted with people who (had) used a sleep product. The first step was to identify a sample of technology-mediated sleep products. The most familiar sleep product is probably the alarm clock. Besides alarms, there are many other sleep-supporting technologies on the market. For example, relaxation-supporting technologies, sleep-measuring apps and devices, and computerized therapies.

2.1.3. Study III: Focus Group Discussions - The Envisioned Sleep Coach

A limitation of the interviews from Study II, as will be discussed in more detail in the Results section, was that they were restricted to existing products, and did not include reflections on what might technically be possible regarding adherence-enhancing strategies. During the interviews, it also proved to be difficult for participants to think of additional functionality that could improve their adherence. To address the limitations

of the interviews, focus groups were organized to discuss adherence-enhancing strategies of a to-be-developed sleep coach. The aim of study III was to gain insight into coachees' attitudes and beliefs toward these adherence-enhancing strategies, for which focus groups are particularly suited (Kuniavsky, 2003).

2.2. Methods

2.2.1. Study I

2.2.1.1. Overview

The meta-analysis was primarily performed to answer the question "How well do coachees adhere to technology-mediated insomnia interventions and diagnostic tools?" and, secondly, to answer the question "How does adherence relate to treatment outcome?" Various databases were queried—Web of Science, Scopus, PubMed, and PsychINFO—on July 8 and 14, 2014, to find studies that investigate insomnia regimes mediated by technology. The used query was: *insomnia* and *Internet-treatment*, *Internet-delivered*,

Internet-based, *Internet-administered*, *Internet intervention*, *computerize*, *online treatment*, *Web application*, *Web-based*, *virtual*, *virtual reality*, *mass media intervention*, *smartphone*, *mobile phone*, *mobile technology*, *text message*, *handheld*, or *PDA* (personal digital assistant). In addition, the references from recent meta-analyses, and systematic reviews on self-help and computerized insomnia therapy (Cheng & Dizon, 2012; Ho et al., 2014; van Straten & Cuijpers, 2009) were screened for potentially relevant publications. Together, this resulted in 448 unique papers of which the abstracts were read and examined (by the first author, CH) for meeting the following exclusion criteria: no main focus on insomnia, no technology involved, treatment that does not include assignments at home, no experiment, or targeted at children. Studies on children were excluded because children's sleep problems often differ from those of adults. Besides, children's bedtimes are partly controlled by the parents. Therefore, interventions targeted at children have other characteristics than interventions for adults and were excluded. A total of 56 papers were read completely and the inclusion of those papers was discussed between the first and second author (CH and JL).

Figure 1 shows the flow diagram for inclusion and exclusion criteria, resulting in 21 papers from which data was retrieved. Due to a lack of reported adherence data in 3 of the papers, only 18 papers were used in the analysis. The papers selected for this meta-analysis can be found in Appendix A.

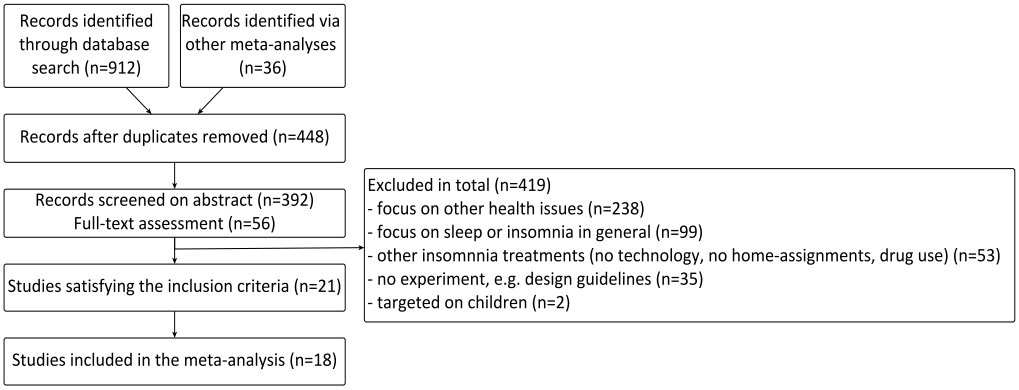


Figure 1. Inclusion and exclusion criteria for papers in the meta-analysis.

2.2.1.2. Description of Included Studies

Of the 18 included studies in this meta-analysis, 12 studies (67%) focused on CBT-I (Table 1 and 2). Oosterhuis and Klip (1997) and Rybarczyk and colleagues (2002) did include most of the CBT-I exercises in their intervention. Out of the 18 studies, 2 (11%) focused on sleep tracking by using an active sleep sampling device. The active sleep sampling device used by Riley and colleagues (Riley, Mihm, Behar, & Morin, 2010) mainly supported sleep restriction and stimulus control. The other standard CBT-I exercises were explained in an additional manual. Lawson and colleagues (2013) used an active sleep sampling device, inspired by Riley's device. They developed an active sleep sampling mobile phone app which focused on sleep tracking, but did not include the other CBT-I components. Lipschitz and colleagues (Lipschitz, Landward, & Nakamura, 2014) also developed a mobile phone app, offering sleep-focused, mind-body bridging exercises. The most important assumption of mind-body bridging for sleep is that the mind needs to be rested to sleep well. Haimov and Shatil (Haimov & Shatil, 2013) studied whether providing cognitive training, such as a memory game, affects sleep.

2.2.2. Study II

2.2.2.1. Participant Selection

In order to establish a purposive sample of users across sleep products, various sleep products were categorized. Based on their background knowledge and a media scan, the authors generated a list of 54 technologies over the course of a few months. This composed list was supplemented with apps because the goal of the Sleepcare project is to design a virtual sleep coach on a mobile phone. The first 25 Android apps and 25 iPhone apps found in Google Play and the iTunes store with the search word "sleep" on November 19, 2012, were added to the product list. A total of 7 apps were unrelated to sleep—3 games, 2 hypnosis apps, 1 unlock, 1 music timer—and were therefore

Table 1. Characteristics of included studies.

	First author	Condition	Number of people	Number of females/ males	Mean age	Sleep problem severity (measure, score)
1	Oosterhuis	Intervention	400	63% female	55	N/A[a]
2	Rybarczak	Intervention	14	22/16	68	PSQI[b], 9.5
		CBT[c]	11			PSQI, 11.9
		Control	13			PSQI, 9.9
3	Ström	Intervention	54	71/38	44	ISI[d], 18.08
		Waiting list	55			ISI, 18.11
4	Suzuki	Intervention	21	16/25	40	N/A
		Waiting list	22			
5	Ritterband	Intervention	22	34/10	N/A	ISI, ≥8
		Waiting list	23			ISI, ≥8
6	Van Straten	Intervention	126	163/84	52	72% rated SQ[e] <6/10
		Waiting list	121			68% rated SQ[e]<6/10
7	Vincent	Intervention	59	79/39	N/A	N/A
		Waiting list	59			
8	Riley	Intervention 1	24	52/38	49	ISI, 8-14 (25 people)
		Intervention 2	33			ISI, 15-21 (53 people)
		SMMT[f]	33			ISI, 22-28 (12 people)
9	Lancee	CCBT-I[g]	216	520/103	52	Sleep-50, ≥19
		CBT-I[h]	202			
		Waiting list	205			
10	Ritterband	Intervention	14	24/4	57	ISI, 17.1
		Waiting list	14			ISI, 15.9
11	Espie	Intervention	55	120/44	49	Met DSM-5[i] criteria
		TAU[j]	54			Met DSM-5 criteria
		IRT[k]	55			Met DSM-5 criteria
12	Haimov	Cognitive training (CogniFit)	34	29/22	72	Met AASM[l] criteria
		Active control[m]	17			Met AASM criteria
13	Lancee	Low depression	198	316/163	47	ISI, 16.73
		Mild depression	182			ISI, 18.63
		High depression	99			ISI, 20.69
						Average ISI, 18.72

	First author	Condition	Number of people	Number of females/ males	Mean age	Sleep problem severity (measure, score)
14	Lancee	With support	129	197/65	48	ISI, 16.95
		Without support	133			ISI, 17.32
15	Lawson	Intervention	36	21/5	34	N/A
16	Van Straten	Intervention	59	83/35	49	PSQI, 12.4
		Waiting list	59			PSQI, 11.7
17	Holmqvist	Intervention	39	55/18	N/A	ISI, 18.72
		CBT-I	34			ISI, 18.50
18	Lipschitz	Intervention	37	27/10	37	ISI, 7.24

[a]Not applicable (N/A).
[b]Pittsburgh Sleep Quality Index (PSQI).
[c]Cognitive behavioral therapy (CBT).
[d]Insomnia Severity Index (ISI).
[e]Sleep quality (SQ).
[f]Self-monitoring minimal treatment (SMMT).
[g]Computerized cognitive behavioral therapy for insomania (CCBT-I).
[h]Cognitive behavioral therapy for insomania (CBT-I).
[i]Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5).
[j]Treatment as usual (TAU).
[k]Imagery relief therapy (placebo) (IRT).
[l]American Academy of Sleep Medicine (AASM).
[m]Active control consisted of word and paint training.

Table 2. Description of included studies.

	First author	Intervention	Delivery	Treatment length	Follow-up length	Post[a]	FU[b]	Adherence measure
1	Oosterhuis	SE[c], SH[d], CTh[e], RX[f]	TV[g]	8 weeks	4.5 months	Q[h]	Q	N/A[i]
2	Rybarczyk	RX, SC[j], SR[k], CTh, SH	Audiotape	6 weeks	4 months	Q&D[l]	Q&D	N/A
3	Ström	CBT-I[m]	Internet	5 weeks	9 months	Q&D	D	N/A
4	Suzuki	CBT-I	Internet	2 weeks	3 weeks	Q	Q	N/A
5	Ritterband	CBT-I	Internet	9 weeks	6 months	Q&D	Q	N/A
6	Van Straten	CBT-I	TV	6 weeks	None	Q&D	N/A	SRep[n]
7	Vincent	CBT-I	Internet	5 weeks	4 weeks	Q&D	Q&D	SRep
8	Riley	ASS[o]/CBT-I	Device	6 weeks	6 weeks	Q&D	Q&D	
9	Lancee	CBT-I	Internet	6 weeks	4 weeks		Q&D	SRep
10	Ritterband	CBT-I	Internet	6-9 weeks	None	Q&D	N/A	Log
11	Espie	CBT-I	Internet	6 weeks	8 weeks	N/A	Q&D	Log
12	Haimov	CTr[p]	PC[q]	8 weeks	None	Q&D	N/A	N/A
13	Lancee	CBT-I	Internet	6 weeks	4 weeks	N/A	Q&D	SRep
14	Lancee	CBT-I	Internet	6 weeks	6 months	Q&D	Q&D	Log
15	Lawson	ASS	App[r]	7 days	None	Q	N/A	Log
16	Van Straten	CBT-I	Internet	6 weeks	3 months	Q&D	Q&D	Log
17	Holmqvist	CBT-I	Internet	6 weeks	8 weeks	Q&D	Q&D	N/A
18	Lipschitz	MBB[s]	Internet	3 days	1 week	Q	Q	SRep

[a] Postintervention measurement instrument.
[b] Follow-up measurement instrument.
[c] Sleep education (SE).
[d] Sleep hygiene (SH).
[e] Cognitive therapy (CTh).
[f] Relaxation (RX).
[g] Television (TV).
[h] Questionnaire (Q).
[i] Not applicable (N/A).
[j] Stimulus control (SC).
[k] Sleep restriction (SR).
[l] Sleep diary (D).
[m] Cognitive behavioural therapy for insomnia (CBT-I).
[n] Self-report
[o] Active sleep sampling (ASS) device.
[p] Cognitive training (CTr) (CogniFit).
[q] Personal computer (PC).
[r] Mobile phone app (app).
[s] Mind-body bridging (MBB).

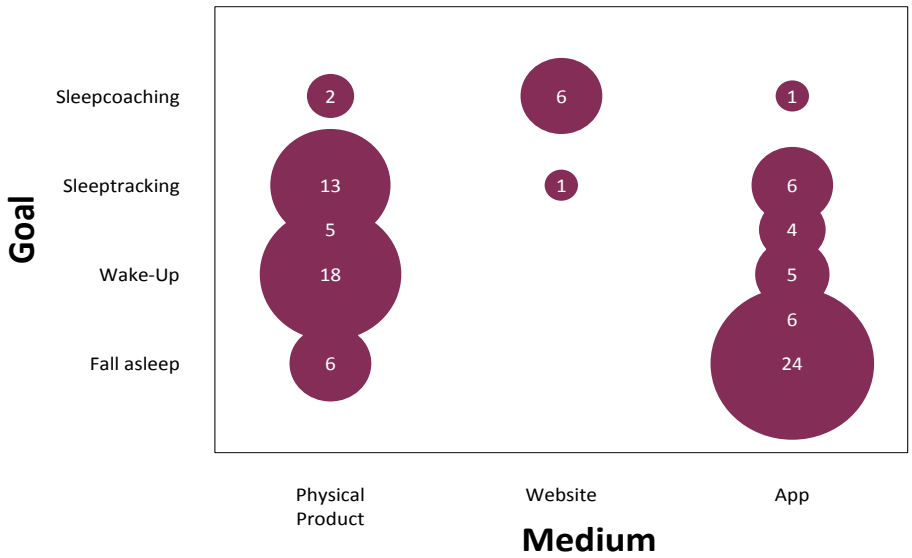


Figure 2. A graph showing the relationship between the goal of sleep products and the medium used. The size of each bubble indicates how many products of the 97 identified sleep products belong to that category.

discarded, resulting in a list of 97 sleep products. The categorization made in this paper aims to be simple and objective. Sleep products were categorized based on their goal and the medium used. Figure 2 shows the distribution of the products across the two dimensions: goal and medium. The size of the bubbles shows how many products belong to the intersections of the categories.

After identifying the categories of existing sleep products, the next step was to learn more about the users' usage and adherence to the sleep products. Interviews were conducted with people who used a sleep product in each of the largest product-medium combinations (e.g., apps that help people fall asleep).

2.2.2.2. Participants

People registered as participants at the Sleepcare project website (www.ikgalekkerslapen.nl) were invited to participate in the interviews if they had ever used a technology-mediated sleep product. In addition, two sleep therapists were asked to invite people who used sleep coaching products, as none of the respondents to the call used a sleep coaching product. A total of 15 Dutch persons agreed to be interviewed—6 (40%) females and 9 (60%) males—their ages ranging from 22 to 65 years (mean 37.5, *SD* 14.8). The mean Pittsburgh Sleep Quality Index (PSQI) (Buysse, Ancoli-Israel, Edinger, Lichstein, & Morin, 2006) score was 8.0 (*SD* = 4.0), with 12 out of the 15 (80%) interviewees having a score above 5, which is the threshold for poor sleep quality classification.

2.2.2.3. Interviews

Besides adherence, the interviews covered other topics to gain insight regarding users' experiences with sleep products. Therefore, the semistructured interviews included both adherence-related questions and questions regarding the factors of the Unified Theory of Acceptance and Use of Technology (UTAUT) model (Venkatesh, Morris, Davis, & Davis, 2003). The described results of the interviews in this paper, however, will only include adherence-related topics. The interviews were conducted in person, by Skype, or by telephone by the first author. The audio of the interviews was recorded. The study was approved by the Human Research Ethics Committee of Delft University of Technology.

2.2.2.4. Analyses

The first author performed the data analysis following the phases of thematic analysis as described by Braun and Clarke (2006). The first author familiarized herself with the data (phase 1) by conducting and transcribing the interviews, and reading the transcripts. While reading, the initial codes were generated bottom-up (phase 2). The first author coded the transcriptions and iteratively generated hierarchical codes and themes (phase 3). Short summaries of the codes and themes related to adherence were written down. The first and last author (CH and WB, respectively) discussed these summaries (phase 4) to form three final adherence-related categories (phase 5). In addition, an independent researcher applied the coding scheme to one of the interviews in order to minimize the threats to confirmability (known as objectivity in quantitative research). The independent coder confirmed the applicability and usefulness of the codes.

2.2.3. Study III

2.2.3.1. Overview

The envisioned coach would use different adherence-enhancing strategies during the entire coaching process. For example, different roles (e.g., motivator and educator) could be played by different virtual characters to increase the effect of the to-be-developed sleep coach (i.e., split-persona effect) (Baylor & Ebbers, 2003). Around 25 strategies were allocated to the coach ranging from strategies involving others (e.g., peers or family members), helping with planning (e.g., setting goals and making commitments), and gaming strategies (e.g., earning points and taking a quiz). These adherence-enhancing strategies were scripted explicitly in the scenarios in order to discuss them in the focus groups.

Table 3. Demographics of the participants per focus group.

Focus groups	Participants, *n* (% female)	Age in years, mean, (*SD*)	Number of participants with a PSQI[a]>5	Expertise
Potential users 1	8 (38)	35 (12)	3 (38)	N/A[b]
Potential users 2	7 (71)	48 (9)	5 (71)	N/A
Coaches 1	4 (75)	51 (8)	N/A	4 coaches (relationships, lifestyle, didactical)
Coaches 2	5 (80)	50(6)	N/A	4 coaches (lifestyle, career),1 psychologist
Sleep experts 1	3 (67)	50 (18)	N/A	1 psychologist, 1 therapist, 1 doctor
Sleep experts 2	4 (75)	47 (14)	N/A	3 researchers, 1 psychologist

[a]Pittsburgh Sleep Quality Index (PSQI).
[b]Not applicable (N/A).

2.2.3.2. Materials

A total of 12 scenarios and 72 claims (see Appendices D and E) were written to evaluate the adherence-enhancing strategies. Scenarios consisted of stories about people, their activities, goals, and motivations regarding a system (Carroll, 2000). Claims stated important design decisions (e.g., about the adherence-enhancing strategies) that needed to be evaluated in the focus group. Furthermore, three fictitious people varying in age, gender, family situation, and readiness-to-change were created to act in the scenarios (see Appendix F for these personas).

2.2.3.3. Procedure and Participants

The scenarios and claims were discussed in six focus groups to evaluate the adherence-enhancing strategies. Two groups consisted of potential users, two groups consisted of coaches, and a further two groups consisted of sleep experts. Demographics of the Dutch participants can be found in Table 3. Each session lasted 2 hours and included a general introduction, an introduction round of the participants, and approximately four animated videos that represented the different scenarios. After watching one video, the participants were asked to individually rate their agreement with the claims on a 7-point Likert scale. Subsequently, participants were asked in turn to react to the claims and discuss their ideas. The sessions were videotaped for later analysis. The study was approved by the Human Research Ethics Committee of Delft University of Technology.

2.2.3.4. Analyses

The analysis was an iterative process of developing codes and themes in line with thematic analysis (Braun & Clarke, 2006). For that, the videotapes of the sessions were transcribed and summarized by the first author. During that recapitulation, several codes emerged and an initial coding scheme of 12 codes was created. The first author coded the summaries according to this scheme. Additionally, a second coder, independent of the project, coded a sample of the summaries — 48 of the 86 claims (56%). The second coder suggested eight additional codes. The two coders came together to discuss the coding scheme and agreed on a new scheme of 15 codes. The coding was improved (with this new scheme) by both coders, and within the sample a Cohen's kappa of .80 was reached. Next, the first author wrote short resumes per theme, making use of quotes.

2.3. Results

2.3.1. Study I

All analyses were completed with the Comprehensive Meta-Analysis statistical package, version 3, and were based on the random-effects model. In the analyses, a distinction was made between experimental compliance and treatment adherence. All studies reported experimental compliance, and most of them (10) also reported treatment adherence (see Table 1 and 2). Experimental compliance was typically determined based on the completion of questionnaires and sleep diaries that were part of the study protocol, for more information see Appendix B. Immediately after the intervention (i.e., postmeasures), the experimental compliance for questionnaires was 78% (95% *CI*:70%-85%), and for sleep diaries 71% (95% *CI*: 65%-77%). At the follow-up assessments, experimental compliance to questionnaires was 72% (95% *CI*: 69%-76%), while for diaries it was 58% (95% *CI*: 52%-64%). These aggregated numbers are displayed in Figure 3. In appendix C, individual numbers per study, aggregated rates, heterogeneity statistics, and publication bias tests can be found. Generally, the analyses indicated a substantial heterogeneity in the data, which supports the choice for a random-effects model. The shapes of the funnel plots and the Egger test did not suggest a significant publication bias.

Treatment adherence was reported in various ways, which can roughly be classified into two groups, namely self-reports and logs. Self-reports refer to questions in which participants were asked how well they adhered to the exercises. The five studies that used self-reports found that 41% (95% *CI*: 36%-46%) of the participants met the adherence criteria set in that study. Logs refer to reports that show how many sessions were completed. Five studies used logs and found that 64% (95% *CI*: 44%-79%) of the

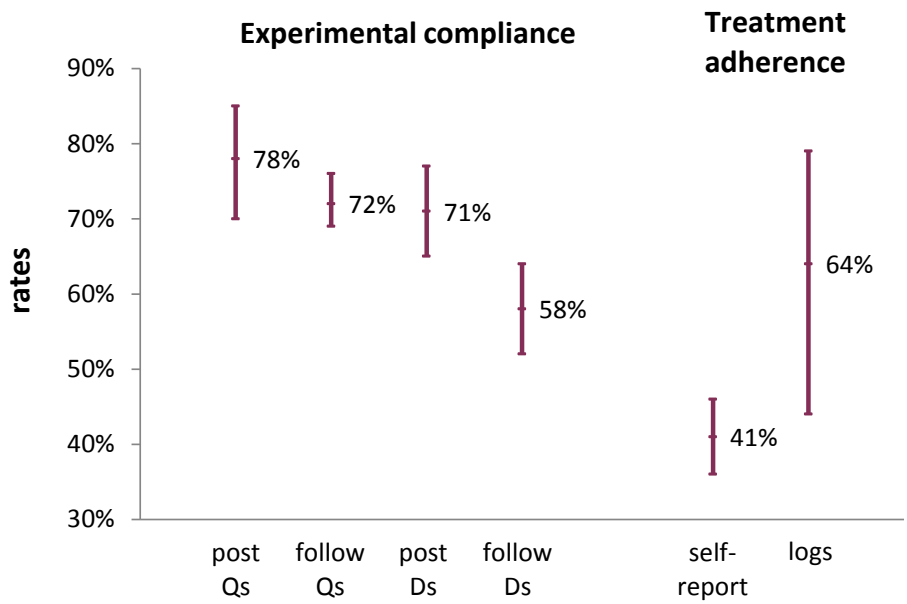


Figure 3. Mean compliance and adherence rates and their 95% CIs. post: posttreatment measurement; follow: follow-up measurement; Qs: questionnaires; Ds: diaries; self-report: self-reported adherence with questions; logs: automatically logged behavior.

participants completed all sessions. If these two kinds of measures are taken together, an average treatment adherence of 52% (95% *CI*: 43%-61%) is reached with reported adherence ranging from 28% (Lawson et al., 2013) to 100% (Lipschitz et al., 2014) across the 10 studies.

In Figure 3 the compliance and adherence rates and their 95% confidence intervals are shown; as can be seen, the self-reported treatment adherence is significantly different from the experimental compliance rates (nonoverlapping confidence intervals). Furthermore, two meta-regressions were run with studies that reported both experimental compliance and treatment adherence in order to discover a possible relationship between these two measures (experimental compliance and treatment adherence): one meta-regression with experimental compliance to postquestionnaires as the explanatory variable and logged treatment adherence as the outcome variable, and the other meta-regression with experimental compliance to follow-up questionnaires as the explanatory variable and self-reported treatment adherence as the outcome variable. These variables were chosen because most data were available for these combinations of variables. Both analyses did not reveal significant relationships between experimental compliance and treatment adherence (both had $p > .05$).

Table 4. Statistics of the meta-regression of adherence and effect size of the individual treatments.

Statistics meta-regression	Coefficient	Standard error	95% CI	Z	*p* (2-sided)
Intercept	0.74	0.20	0.35-1.13	3.69	<.001
Adherence	0.79	0.35	0.10-1.47	2.25	.03

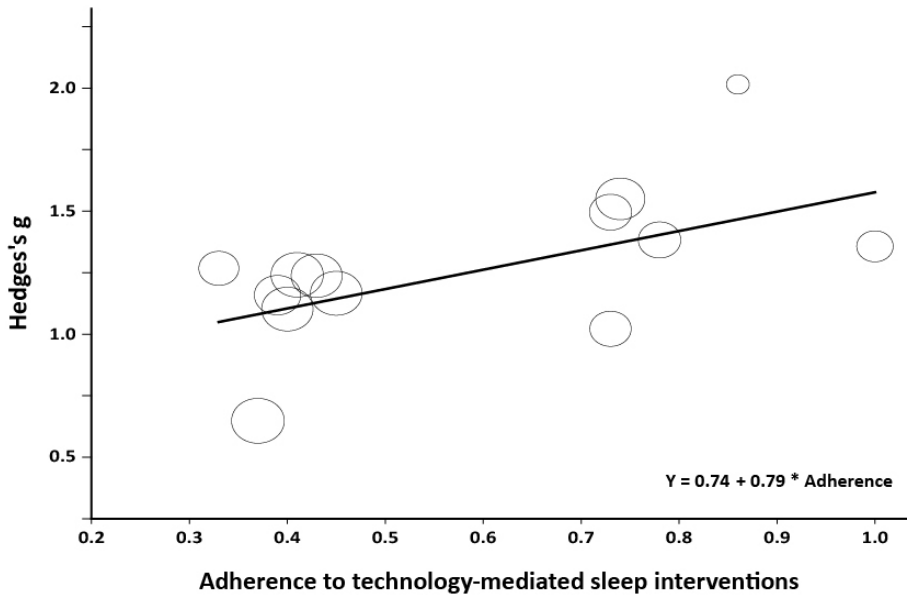


Figure 4. Meta-regression of adherence on effect size of treatments. The circles represent the individual studies. The circle size indicates the weight of the study. The effect size is given in standard difference in means.

Lastly, the relationship between treatment adherence and the effect size of the individual treatments was explored. Figure 4 and Table 4 show the results of the meta-regression analysis. The analysis revealed a significant model (Q_{model}= 5.05, df = 1, p = .03), with a coefficient of 0.79 (Z = 2.25, p = .03) for adherence. In other words, treatment adherence and treatment effect are positively correlated. For example, if adherence increases with 0.30 (30%), this would coincide with a 0.24 increase in effect size (Hedge's g) of the treatment, which is an increase of a small effect size of 0.20. The analysis also found that 75.4% (I^2 = 75.4, Q = 48.87, df = 12) of the total variance in effect size could be explained by the variation between the studies. Of this 75.4%, 40% (R^2 = .407, T^2_{total} = .059, $T^2_{unexplained}$ = .035) could be explained by treatment adherence.

Usage

- Intention
  - Two reasons for usage
    - Overcome sleep problems
    - Interest in the product
- When used
  - Sleep trackers, alarms, relaxation: used in the evenings
  - Sleep coaches: varying usage times

Effectiveness

- Per product type
  - Therapy-related products: no noticeable effect
  - Alarms: ambiguous effect, wake-up is okay, but not waking up better

Adherence

- Keep using
  - Therapy-related products: personal attitude
  - Consumer products: need functionality
- Not using
  - Consumer products
    - No need for functionality (anymore)
    - Product does not work
    - Forget to use product

Box 1. Main themes mentioned by participants in the interviews.

2.3.2. Study II

2.3.2.1. Overview

The three main categories related to adherence are usage, effectiveness, and adherence (see Box 1).

2.3.2.2. Usage

Two initial reasons for using a product emerged from the interviews. First, all interviewees used the product to overcome some of their sleeping troubles. Interviewees wanted to wake up better, initiate or maintain sleep, and/or increase insight into their sleep. Second, some interviewees used a product because they thought the product in itself was interesting. Above all, this holds for the sleep-tracking apps. Most products—

alarms, automatic sleep trackers, and relaxation support—were used in the evening before going to sleep. Most of those products, however, were not used on the weekend. Sleep coach usage varied, depending on the kind of assignments included in the product (e.g., diary, relaxation, sleep hygiene exercise, bedtime scheduling).

Participants' quotes regarding reasons to start using sleep products were as follows (translated from Dutch):

> *The reason was that in my opinion I was awake too often, and too long. I could not fall asleep anymore.* [Interview #11, online sleep therapy]
>
> *Friends of mine had the app and I wanted to try it as well.* [Interview #4, sleep tracking app]
>
> Participants' quotes regarding usage of sleep products were as follows (translated from Dutch):
>
> *I turn it on in the evening when I am lying in bed and want to go to sleep*. [Interview #1, relaxation app]
>
> *Every day I needed to get up I used it, but on the weekends, for example, when I don't need to get up I didn't do anything with the app.* [Interview #7, sleep tracking app]
>
> *Actually, I did it whenever it suited me [about filling in a sleep diary].* [Interview #10, online sleep therapy]

2.3.2.3. Effectiveness

One of the initial arguments for using a product was to overcome some kind of sleeping problem. However, the online sleep therapies were not perceived as having an effect on the interviewees' sleep problems. Additionally, interviewees mentioned that it was hard to determine if the therapy improved sleep in the long term because they tried several things. Nevertheless, most interviewees took some advice that worked for them and continued applying it. Furthermore, sleep tracking apps as well as online sleep coaches provided the interviewees with more insight into their sleep and habits. Products that wake interviewees up (i.e., smart alarm apps and wake-up lights) were assessed ambiguously. Both types of products did what was expected of them, namely wake the interviewee up. However, the effect of waking up better with the product was doubted. Moreover, smart alarms did not seem to fit into interviewees' daily lives (see quote below from interview #4, sleep tracking app).

Participants' quotes regarding the effectiveness of sleep products were as follows (translated from Dutch):

> *The goal to sleep better was not reached*. [Interview #12, online sleep intervention]

I have no clue if it helped, because it is going better at the moment, but I did other things in that same time period. [Interview #10, online sleep therapy]

Also getting out of bed when I am awake for more than 30 minutes. The advice has helped, yes. [Interview #10, online sleep therapy]

It measures the sleep debt that you are building up, that was effective. [Interview #6, sleep tracking app]

It did what it supposed to do, wake me up. [Interview #7, sleep tracking app]

Still not very well, but it became a little bit better, a little bit more pleasant [about waking up with a wake-up light]. [Interview #5, sunrise alarm]

Problem of the app [smart alarm] is that you do not know what time you will wake up exactly. If I have an appointment somewhere I need an hour to get ready. If you do not know how late your alarm will go, it is hard to plan. [Interview #4, sleep tracking app]

2.3.2.4. Adherence

In general, interviewees perceived their own usage as sufficient. Interviewees especially perceived their own personal attitude, beliefs, and willpower as important for adherence. These personal characteristics were regarded as particularly important for adherence to therapy-related products. The usage of consumer products (e.g., an alarm clock) was continued, because the interviewees needed the functionality. The main arguments for not using a consumer product were (1) no perceived need for the product, (2) a perceived lack of effectiveness, and (3) the interviewee forgot to use the product.

Participants' quotes regarding the satisfaction about adherence of sleep products were as follows (translated from Dutch):

It went well. I cannot remember not doing the exercises. [Interview #13, online sleep therapy]

[About doing the exercises everyday] Well, that went ok. [Interview #10, online sleep therapy]

I use it 3 or 4 times a week, depending on my needs. [Interview #1, relaxation app]

Participants' quotes regarding the effect of personal attitude, beliefs and willpower on adherence of sleep products were as follows (translated from Dutch):

I tried to keep myself to it as much as possible, and of course I missed a day now and then, but I tried really hard. [Interview #10, online sleep therapy]

You cannot just resign and accept your sleep problem. [Interview #10, online sleep therapy]

> *I was really motivated, so that makes a difference.* [Interview #10, online sleep therapy]
>
> *[What dragged you through it?] My will. I intended to do it. I started it and I wanted to get a grip on my sleep problem, so I had to follow through. [So your own determination?] Yes, without discipline you will not succeed.* [Interview #11, online sleep therapy]
>
> *I felt like, I started it, so I should finish it.* [Interview #12, online sleep intervention]
>
> *You have to be serious about it. It is a therapy that you really have to complete, otherwise it will not have an effect. So, you have to believe in it.* [Interview #13, online sleep therapy]
>
> *If you do not recognize the need to change, you should not start it.* [Interview #13, online sleep therapy]

Participants' quotes regarding reasons for using and not using sleep products were as follows (translated from Dutch):

> *You have to set an alarm, anyway.* [Interview #6, sleep tracking app]
>
> *During the holidays there is no need for an alarm.* [Interview #5, sunrise alarm]
>
> *I did not have the impression that the app could change my sleeping pattern.* [Interview #6, sleep tracking app]
>
> *I simply forgot it.* [Interview #4, sleep tracking app]

2.3.3. Study III

2.3.3.1. Overview

The most obvious emerging themes in the focus groups were *users in control* and *doing it for your own sake*. In general, participants believed in the personal strengths and willpower of users to adhere to the proposed sleep coach. Furthermore, the adherence-enhancing strategies and motivation were discussed. See Box 2 for an overview of the results.

2.3.3.2. Users in Control

Potential users, coaches, and sleep experts agreed that the users should be in control. Different arguments were given. The coaches and sleep experts mainly argued that giving the user more control increases commitment and motivation. The potential users argued that they use the sleep coach for their own sake, so they want to be in control themselves. Another argument was that not being in control could lead to irritation. Aspects that participants believed the users should be able to control were the following: reminders, amount of information given by the app, scheduling exercises, decisions about motivation level, sharing therapy progress, sharing the outcome of questionnaires, and parameters shown in sleep diary overview.

Users being in control

- *Control increases commitment and motivation*

Doing it for own sake

(phrase that was strongly believed in was "I do it for my own sake")

- *Motivation: three conflicting ideas*
  - If coach is downloaded, then the user is motivated
  - Downloading does not imply motivated usage
  - Motivation can arise while using
- *Adherence-enhancing strategies*
  - Awarding points for progress
    - Not seen as appropriate for sleep coach, however, awarding points can work against own expectations
  - Giving compliments
    - Not too often, not for nonsignificant actions
    - Should contain context, and vary over time
  - Providing reminders
    - Should not be necessary, however, they are practical
    - Reminders are perceived as positive when set by the users
  - Provide rationale: two types of people
    - Type 1: first experience exercise, then explanation
    - Type 2: first explanation, then perform exercise
  - I am not the only one
    - Provide a forum, stories from others, amount of app users, statistics

Box 2. Main themes mentioned in the focus groups with potential users, coaches, and sleep experts.

2.3.3.3. Doing it for Your Own Sake

The other interesting theme was *doing it for your own sake*. In one scenario, there was an example exercise which entails making a list of people who can help you. In general, this exercise met resistance by the potential users. The idea that you have to solve your problems yourself was dominant. Users would feel ashamed to ask for support, and they believed the virtual sleep coach should help them. On the other hand, the coaches stated that thinking about social resources, such as family and peers, could really help people. The coaches mentioned that coachees usually consult a coach exactly because they try to solve their problems themselves, instead of asking their social resources for help. One potential user shared that only informing other people

about her sleep problems and therapy already helped her a lot, even without asking for support. Nonetheless, the general mind-set was that people use such a coach for their own sake, and that they are and should be able to take responsibility for their own adherence.

2.3.3.4. Motivation

The claims underlying the envisioned usage scenarios stated that users should be motivated before they start sleep treatment, otherwise the probability of dropping out would be too high. The focus groups with the sleep experts manifested three different ideas about motivation. Some of the sleep experts argued that people will be motivated at least a little bit when they have downloaded the app, since that requires some effort. On the other hand, it was also argued that someone could show interest in the sleep coach, but he or she would not necessarily be motivated to use the sleep coach. Third, it was argued that motivation could arise during different phases of a therapy; for example, after someone performs an exercise and experiences its effects. In that situation, users would not need to be highly motivated at the beginning of the therapy.

2.3.3.5. Adherence-Enhancing Strategies

Several adherence-enhancing strategies and ideas to increase motivation were scripted in the scenarios (e.g., awarding points, compliments, reminders) and are discussed below.

Awarding Points for Progress

In general, participants reacted adversely to the idea of awarding points as described in the scenarios, mainly because the sleep coach was seen as a serious program for adults. Furthermore, it was believed that a point system is not appropriate for sleep exercises, but more for workout programs. Nevertheless, a few participants spoke up and said that they liked the idea of points. A few stories came up about how awarded points motivated participants in other domains against their own expectations. Thus, points might improve adherence, despite users' initial reluctance.

Compliments

Furthermore, both the coaches and the potential users made negative remarks about the compliments. In principle, both groups thought compliments could enhance a user's experience, but compliments should not be given too often, or for nonsignificant actions. They argued that compliments should contain context and should vary over time. Otherwise, compliments would not increase motivation.

Reminders

Reminders were embraced, as long as users are in control of those reminders. The users wanted to set the reminders themselves, because sometimes "you just forget to do something." On the other hand, some users stated they do not need reminders, since they are using the sleep coach for their own sake. Besides that, they argued that they are

adults, are motivated, and have self-discipline. Both the coaches and the sleep experts agreed with those potential users and thought that reminders should not be necessary. However, from a practical point of view, they understood that people sometimes do forget to do therapy exercises.

Ideas Generated by the Participants

Other ideas to improve motivation mentioned by the participants were as follows: provide a rationale, show statistics, decrease the feeling of being alone, positive feedback, taking small steps, choosing your own coach, demanding a small investment before starting, and showing how much effort users have already invested.

According to the sleep experts, rationales for doing an exercise should be given before users start an exercise. However, the potential users and coaches mentioned there are two types of people: people who want to know how and why things are the way they are, and people who just want to experience an exercise and afterward gain an understanding of that exercise.

Secondly, different ideas were offered to ensure that users do not feel as if they are the only ones suffering from sleep problems. Ideas included a forum (suggested by users and coaches), reading stories from peers (suggested by coaches), and a measure that indicates how many people are using the app (suggested by sleep experts). The idea was that decreasing the feeling of being the only one with sleep problems could increase the motivation of users to adhere to the sleep therapy.

2.4. Discussion

2.4.1. Study I

The meta-analysis of adherence rates found a mean experimental compliance of at least 70%, except for the follow-up diaries. Filling out a diary every day for a full week a few months after the intervention requires quite some effort, which might explain a lower adherence rate (58%) to follow-up diaries than to the other experimental compliance measures. The average self-reported treatment adherence was 41%, whereas the average logged adherence was higher at 64%. This is surprising because the self-reported adherence was less “strict” than the logged adherence; for instance, users were categorized as adherent when they reported doing an exercise a certain number of times (e.g., more than 4 times a week), while the logged adherence rate was based on doing all exercises. The average treatment adherence rate (logged and self-reported, combined) was 52%. Although self-reports and logs are not exactly the same, they both measure adherence and are similar enough to be combined. Nevertheless, this general adherence rate of 52% should be interpreted carefully.

Furthermore, this meta-analysis confirmed that treatment adherence is positively related to treatment effect when it comes to technology-mediated insomnia treatment. Moreover, this analysis showed that experimental compliance and treatment adherence are not related. In other words, the percentage of participants who filled out questionnaires after the intervention was not found to be an indication of how well people adhered to the treatment. Therefore, it seems important to distinguish between experimental compliance and treatment adherence.

The quality of the individual studies was not assessed using a predefined algorithm, which might be a limitation. However, the included studies were all published in peer-reviewed journals and proceedings, which warrant an acceptable level of quality. Besides, Glass and colleagues argue that all studies should be included (Glass, MacGaw, & Smith, 1984). According to them, all studies should be reviewed in context with each other regarding the topics at issue, not necessarily regarding the overall quality of each study. Since adherence is the main focus of this paper, instead of examining a possible relationship between general study quality and adherence (Glass et al., 1984), the methodological differences of measuring adherence were reviewed by differentiating between experimental compliance and treatment adherence, and self-reported and logged adherence.

2.4.2. Study II

The aim of the interviews was to gain more insight into the reasons why coachees adhere to technology-mediated sleep products. Surprisingly, interviewees were quite satisfied with their own usage, which departs from the average 52% adherence rate found in the meta-analysis. The reasons why people started using a product were either out of interest or to overcome sleep problems. However, the products' effectiveness was doubted by the interviewees and was given as a reason to stop using a product. In interviewees' own opinions, they continued to use consumer products because they needed the functionality, whereas they adhered to therapy-related products because of their own attitudes, beliefs, and willpower. Previous research has also identified functionality as an important determinant for adherence in online sleep treatment (Middlemass et al., 2012). Reasons for nonadherence were as follows: no need for the functionality, lack of effectiveness, or just forgetfulness.

Furthermore, it seemed challenging for interviewees to identify adherence-enhancing strategies in the products. It was also difficult for them to come up with an answer to the question of what could be added to the product to help them continue to use the product.

2.4.3. Study III

Focus groups were organized to discuss adherence-enhancing strategies. In addition to motivation, *users in control* and the awareness to *do it for your own sake* proved to be important for adherence. The focus groups provided insights in the up- and downsides of adherence strategies, such as awarding points, compliments, reminders, and community building.

2.5. General Discussion

2.5.1. Positive attitudes towards adherence

The interviews and focus groups both revealed that people strongly believe willpower is an effective adherence strategy. Participants believed that their personal attitudes, beliefs, and motivation would ensure that they stick to their intentions of using a product. This result should be interpreted with caution because of three phenomena. First of all, sleep deprivation increases ego depletion (Barnes, Schaubroeck, Huth, & Ghumman, 2011). In other words, when people are tired their willpower decreases and it will become more difficult to adhere to anything, including a virtual sleep coach. Second, the interviewees attributed their adherence to their own commitment and attitude, while nonadherence was attributed to malfunctioning of the product. This result should also be interpreted with caution because this phenomenon is in accordance with the self-serving bias. The self-serving bias states that successes are attributed to internal factors, while failure is attributed to external factors (Zuckerman, 1979). Therefore, the "good" adherence rates in the interviews were attributed to the interviewees' own willpower. Third, the participants in the focus groups were quite optimistic about their anticipated future adherence. Being optimistic about oneself and the future is one of the most robust biases (optimism bias) in psychology (Buehler, Griffin, & Ross, 1994; Weinstein, 1980). Several explanations for this unrealistic optimism has been offered, for example, ignoring everything that could go wrong (Buehler et al., 1994), putting too much weight on current intentions (Koehler & Poon, 2006), or having too much faith in willpower for future events (Helzer & Gilovich, 2012). These three phenomena provide reasons for treating participants' optimism toward adherence with caution.

2.5.2. Aversion to adherence-enhancing strategies

Apart from relying on willpower for adherence, aversion to adherence-enhancing strategies emerged during the focus groups. Therefore, when designers implement adherence-enhancing strategies they should not assume that users would initially agree with the usefulness of these strategies.

Various design principles for a virtual sleep coach can be adopted from the interviews and focus groups. The first design principle covers functionality. During the first usage phase, the sleep coach should immediately tickle users' interest, for example, by providing automatic sleep tracking. In the interviews, it appeared that interest made coachees start using products. Next, the sleep coach can provide an already needed functionality (e.g., an alarm clock). According to the interviews, a needed functionality ensures that users keep using a product. Lastly, reminders need to be a part of the sleep coach. Reminders make sure that users do not simply forget to adhere to the coach. Both the participants in the interviews and focus groups indicated that sometimes they just forget to use a product. Participants in the focus groups showed a positive attitude toward reminders as long as the users were in control over the reminders. Therefore, including reminders in a sleep coach would be a good first step in future research to increase adherence.

A second design principle could be to withhold adherence support at the start of the intervention (i.e., to postpone possible help by a virtual sleep coach). In this way, the coachees are acknowledged and respected as serious, motivated, and autonomous adults. Coachees can prove that they adhere to the assignments of the sleep coach; however, the virtual coach can detect when coachees fail to do their assignments, and then offer support. This support can take different forms (reminders, compliments, awarding points, etc.) and can be varied over time based on the needs of the coachee.

A third design principle that can be applied is explaining why willpower does not guarantee success. After such an explanation, the understanding of the added value and acceptance of adherence-enhancing strategies might increase. On top of that, users could be given the control over the employment of adherence-enhancing strategies.

In the authors' opinion, the most important overall design principle is balance. Coachees should not feel overwhelmed with adherence-enhancing strategies, but appreciate some occasional support. Personalization of the virtual sleep coach can ensure that the perfect balance is reached for each and every user. For example, some users might need and appreciate reminders for filling out a sleep diary every day, while other users are more likely to forget to do their relaxation exercises.

2.5.3. Measuring adherence

Lastly, we want to stress that studies should measure and report treatment adherence, and make a distinction between experimental compliance and treatment adherence. It is important that future studies measure and report adherence rates, since it is only by the adherence measure that it can be established whether the treatment actually induces the observed outcome. The frequently made statement that adherence is important for the outcome of a treatment (Beun, 2012; Donkin et al., 2011; Eysenbach, 2005) seems to be supported by the findings of the meta-regression between effectiveness and treatment adherence. As a correlational analysis does not provide insight into the direction of a causal relationship, it remains unclear how effectiveness and adherence influence each other. Nevertheless, if coachees do not follow the treatment protocol

(i.e., adherence rates close to zero), the outcome could be attributed to other things outside the intervention (Gould & Clum, 1993), for example, to the waiting-list effect. Furthermore, it is important to make a distinction between experimental compliance and treatment adherence, since these seem to be two distinct constructs as the meta-analysis found no correlations. An earlier meta-analysis about the effectiveness of CCBT-I found a rather good "adherence rate" of 78% (Cheng & Dizon, 2012). However, this rate would be considered as experimental compliance according to the definition used in this paper. Similar experimental compliance rates—79%, 72%, 70%, and 57%—were found by the meta-analysis, although treatment adherence was significantly lower. The average self-reported treatment adherence was 42%, whereas the logged treatment adherence was 64%. Although no significant difference between these two measures was found, it is important to consider how adherence is defined and measured. A study (Stone, Shiffman, Schwartz, Broderick, & Hufford, 2003) that compared a paper diary with an electronic diary found a tremendous difference between self-reported adherence (90.5%) and logged adherence (10.9%) for a paper diary. Lastly, the question remains whether adherence in experimental settings resembles adherence in nonexperimental real-life settings. It could be that adherence rates in experiments are higher than in real-life situations. One possible explanation is the sunk-cost fallacy (Arkes & Blumer, 1985). To illustrate, experiments demand more from participants regarding (pre-) measurements and participants might therefore be more committed to the intervention. When starting a treatment, they have already invested more time (i.e., the sunk cost) compared to patients in nonexperimental settings, and are therefore less likely to drop out.

2.5.4. Research quality

In order to review the quality of our research, it is helpful to know what we did to take care of the credibility, transferability, dependability, and confirmability of our studies (Lincoln & Guba, 1985). Firstly, threats to all four concepts were minimalized by utilizing three different research methods—meta-analysis, interviews, and focus groups. Furthermore, the credibility of our findings is also consolidated by data source triangulation—literature, current users, potential users, coaches, and sleep experts. Additionally, honesty from our informants was reinforced by stating there are no right or wrong answers, and by allowing them the possibility to withdraw at any moment. We also had regular debriefing sessions between the executors and supervisors in order to strengthen credibility. The level of transferability to other application fields can only be judged by the readers, since they have the knowledge of these other domains (Shenton, 2004). Furthermore, future work can be done to replicate these findings in other fields. Transferability and dependability assessments are supported by descriptions of the research methods and Appendices A-F. Lastly, confirmability was addressed by audit trials and the second coders.

2.6. Conclusions

In conclusion, treatment adherence seems important for the effectiveness of technology-mediated insomnia treatments. Individuals expect that they will adhere well to such treatments and would not gain much from adherence-enhancing strategies. They believe willpower is an effective adherence strategy. The 52% average treatment adherence reported in this paper, however, suggests that there is room for improvement. A virtual coach should be able to cope with this "adherence bias," and persuade users to accept adherence-enhancing strategies (e.g., reminders, compliments, and community building). Future research is needed to test the four derived design principles for a virtual coach, which might help to realize a substantial improvement.

2.7. References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed ed.). Arlington, VA: American Psychiatric Publishing.

Ancoli-Israel, S., & Roth, T. (1999). Characteristics of insomnia in the United States: results of the 1991 National Sleep Foundation Survey. I. Sleep, 22, S347 - Corine: Supplement??

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. Organizational Behavior and Human Decision Processes, 35(1), 124-140. doi: 10.1016/0749-5978(85)90049-4

Baglioni, C., Battagliese, G., Feige, B., Spiegelhalder, K., Nissen, C., Voderholzer, U., . . . Riemann, D. (2011). Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. J Affect Disord, 135(1), 10-19.

Barnes, C. M., Schaubroeck, J., Huth, M., & Ghumman, S. (2011). Lack of sleep and unethical conduct. Organizational Behavior and Human Decision Processes, 115(2), 169-180. doi: 10.1016/j.obhdp.2011.01.009

Baylor, A., & Ebbers, S. (2003). The pedagogical agent split-persona effect: When two agents are better than one. Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications.

Benca, R. M. (2005). Diagnosis and treatment of chronic insomnia: a review. Psychiatric services, 56(3), 332-343.

Beun, R. J. (2012). Persuasive strategies in mobile insomnia therapy: alignment, adaptation, and motivational support. Personal and Ubiquitous Computing, 1-9. doi: 10.1007/s00779-012-0586-2

Beun, R. J., Griffioen-Both, F., Ahn, R., Fitrianie, S., & Lancee, J. (2014). Modeling Interaction in Automated E-Coaching-A Case from Insomnia Therapy. Paper presented at the COGNITIVE 2014, The Sixth International Conference on Advanced Cognitive Technologies and Applications.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative research in psychology, 3(2), 77-101. doi: 10.1191/1478088706qp063oa

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the" planning fallacy": Why people underestimate their task completion times. J Pers Soc Psychol, 67(3), 366. doi: 10.1037/0022-3514.67.3.366

Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Lichstein, K. L., & Morin, C. M. (2006). Recommendations for a standard research assessment of insomnia. Sleep: Journal of Sleep and Sleep Disorders Research; Sleep: Journal of Sleep and Sleep Disorders Research.

Carroll, J. M. (2000). Five reasons for scenario-based design. Interacting with computers, 13(1), 43-60.

Cheng, S. K., & Dizon, J. (2012). Computerised cognitive behavioural therapy for insomnia: A systematic review and meta-analysis. Psychotherapy and Psychosomatics, 81(4), 206-216. doi: 10.1159/000335379

Christensen, H., Griffiths, K. M., & Farrer, L. (2009). Adherence in Internet interventions for anxiety and depression: Systematic review. J Med Internet Res, 11(2).

Daley, M., Morin, C. M., LeBlanc, M., Grégoire, J.-P., & Savard, J. (2009). The economic burden of insomnia: direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms, and good sleepers. Sleep, 32(1), 55.

Donkin, L., Christensen, H., Naismith, S. L., Neal, B., Hickie, I. B., & Glozier, N. (2011). A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res, 13(3).

Edinger, J. D., & Means, M. K. (2005). Cognitive-behavioral therapy for primary insomnia. [Review]. Clin Psychol Rev, 25(5), 539-558. doi: 10.1016/j.cpr.2005.04.003

Espie, C., Kyle, S., Williams, C., Ong, J., Douglas, N., Hames, P., & Brown, J. (2012). A randomized, placebo-controlled, trial of online Cognitive Behavioral Therapy for chronic Insomnia Disorder delivered via an automated media-rich web application. Sleep, 35(6), 769.

Eysenbach, G. (2005). The law of attrition. J Med Internet Res, 7(1).

Glass, G. V., MacGaw, B., & Smith, M. L. (1984). Meta-analysis in social research: Sage Beverly Hills, CA.

Gould, R. A., & Clum, G. A. (1993). A meta-analysis of self-help treatment approaches. Clinical Psychology Review, 13(2), 169-186.

Haimov, I., & Shatil, E. (2013). Cognitive Training Improves Sleep Quality and Cognitive Function among Older Adults with Insomnia. Plos One, 8(4). doi: 10.1371/journal.pone.0061390

Hebert, E. A., Vincent, N., Lewycky, S., & Walsh, K. (2010). Attrition and adherence in the online treatment of chronic insomnia. Behavioral Sleep Medicine, 8(3), 141-150. doi: 10.1080/15402002.2010.487457

Helzer, E. G., & Gilovich, T. (2012). Whatever Is Willed Will Be A Temporal Asymmetry in Attributions to Will. Personality and Social Psychology Bulletin, 38(10), 1235-1246.

Ho, F. Y.-Y., Chung, K.-F., Yeung, W.-F., Ng, T. H., Kwan, K.-S., Yung, K.-P., & Cheng, S. K. (2014). Self-help cognitive-behavioral therapy for insomnia: A meta-analysis of randomized controlled trials. Sleep Medicine Reviews.

Horsch, C., Brinkman, W. P., van Eijk, R., & Neerincx, M. (2012). Towards the usage of persuasive strategies in a virtual sleep coach. Paper presented at the UK HCI 2012 Workshop on People, Computers & Psychotherapy, Birmingham.

Koehler, D. J., & Poon, C. S. (2006). Self-predictions overweight strength of current intentions. Journal of Experimental Social Psychology, 42(4), 517-524.

Kuniavsky, M. (2003). Observing the user experience. San Francisco: Morgan Kaufmann Publishers.

Lawson, S., Jamison-Powell, S., Garbett, A., Linehan, C., Kucharczyk, E., Verbaan, S., . . . Morgan, K. (2013). Validating a mobile phone application for the everyday, unobtrusive, objective measurement of sleep. Paper presented at the Conference on Human Factors in Computing Systems - Proceedings.

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry (Vol. 75): Sage.

Lipschitz, D. L., Landward, R., & Nakamura, Y. (2014). An exploratory study of an online mind-body program for poor sleepers in a community sample. European Journal of Integrative Medicine, 6(1), 48-55. doi: 10.1016/j.eujim.2013.10.002

Middlemass, J., Davy, Z., Cavanagh, K., Linehan, C., Morgan, K., Lawson, S., & Siriwardena, A. N. (2012). Integrating online communities and social networks with computerised treatment for insomnia: a qualitative study. British Journal of General Practice, 62(605), e840-e850.

Morin, C., LeBlanc, M., Daley, M., Gregoire, J., & Merette, C. (2006). Epidemiology of insomnia: prevalence, self-help treatments, consultations, and determinants of help-seeking behaviors. Sleep medicine, 7(2), 123-130.

Morin, C. M. (1993). Insomnia: Psychological assessment and management: Guilford Press New York.

Morin, C. M., Beaulieu-Bonneau, S., LeBlanc, M., & Savard, J. (2005). Self-help treatment for insomnia: a randomized controlled trial. Sleep, 28(10), 1319-1327.

Morin, C. M., & Espie, C. A. (2003). Insomnia: A clinician's guide to assessment and treatment (Vol. 1): Springer.

Morin, C. M., Gaulier, B., Barry, T., & Kowatch, R. A. (1992). Patients' acceptance of psychological and pharmacological therapies for insomnia. Sleep: Journal of Sleep Research & Sleep Medicine.

Ohayon, M. M. (2002). Epidemiology of insomnia: what we know and what we still need to learn. Sleep medicine reviews, 6(2), 97-111.

Oosterhuis, A., & Klip, E. C. (1997). The treatment of insomnia through mass media, the results of a televised behavioural training programme. Social Science & Medicine, 45(8), 1223-1229. doi: 10.1016/s0277-9536(97)00041-5

Perlis, M. L., Smith, M. T., Cacialli, D., Nowakowski, S., & Orff, H. (2003). On the comparability of pharmacotherapy and behavior therapy for chronic insomnia: commentary and implications. Journal of psychosomatic research, 54(1), 51-59.

Riley, W. T., Mihm, P., Behar, A., & Morin, C. M. (2010). A computer device to deliver behavioral interventions for insomnia. Behavioral Sleep Medicine, 8(1), 2-15. doi: 10.1080/15402000903425314

Rosekind, M. R., & Gregory, K. B. (2010). Insomnia risks and costs: health, safety, and quality of life. Am J Manag Care, 16(8), 617-626.

Rybarczyk, B., Lopez, M., Benson, R., Alsten, C., & Stepanski, E. (2002). Efficacy of two behavioral treatment programs for comorbid geriatric insomnia. Psychology and aging, 17(2), 288.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. Education for information, 22(2), 63-75.

Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., & Hufford, M. R. (2003). Patient compliance with paper and electronic diaries. Controlled clinical trials, 24(2), 182.

van Straten, A., & Cuijpers, P. (2009). Self-help therapy for insomnia: a meta-analysis. Sleep Medicine Reviews, 13(1), 61-71.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.

Vermeire, E., Hearnshaw, H., Van Royen, P., & Denekens, J. (2001). Patient adherence to treatment: three decades of research. A comprehensive review. Journal of clinical pharmacy and therapeutics, 26(5), 331-342.

Vincent, N., & Lionberg, C. (2001). Treatment preference and patient satisfaction in chronic insomnia. Sleep.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. J Pers Soc Psychol, 39(5), 806. doi: 10.1037/0022-3514.39.5.806

World Health Organization, W. (2003). Adherence to long-term therapies: evidence for action. Geneva, Switzerland: World Health Organization.

Zuckerman, M. (1979). Attribution of success and failure revisited, or: The motivational bias is alive and well in attribution theory. Journal of personality, 47(2), 245-287. doi: 10.1111/j.1467-6494.1979.tb00202.x

3. The Downward Spiral of Ambivalence

Doubts about Insomnia, Cognitive Behavioural Therapy, and a Pre-treatment Motivation Module

Abstract

Background - The trans-theoretical model describes six stages people advance through when changing behaviour. Interventions that match a person's stage are believed to be more effective and to improve adherence to the intervention.

Method - A pre-treatment motivation module for insomnia therapy was developed to test if people can be supported to progress to the most ideal stage for starting the treatment. The pre-treatment motivation module was a self-help workbook based on motivational interviewing.

Results - In this uncontrolled study, thirty-three participants filled out a baseline questionnaire. At baseline, most participants were at the contemplation stage. At the post-test, the results showed that only two participants advanced to the next stage, nine remained in the same stage, and four regressed. However, it should be noted that 12 participants (52%) did use the pre-treatment module less often as the prescribed two times. In addition, qualitative data revealed two types of ambivalence: (1) people's ambivalence about behaviour change; (2) people's ambivalence about the pre-treatment motivation module.

Discussion and conclusion - This ambivalence could have led to the low adherence rates and may be explained by the decision theory about expected value of sample information. We expect that designers of other behaviour change interventions could benefit from the insight that people are not only ambivalent about the behaviour change, but also about interventions and tools that support the behaviour change process.

Chapter will be submitted as:

Horsch, C. H. G., Lancee, J., Beun, R. J., Neerincx, M. A., & Brinkman, W. P. (submitted). *The Downward Spiral of Ambivalence. Doubts about Insomnia, Cognitive Behavioural Therapy, and a Pre-treatment Motivation Module*. Paper presented at the e-Coaching for Health and Wellbeing 2017, Amsterdam.

3.1. Introduction

Approximately 10% of the population suffers from insomnia (Ohayon, 2002), meaning they have problems with falling asleep or staying asleep and suffer from daytime consequences (American Psychiatric Association, 2013). Insomnia is associated with several problems, for example, lack of concentration, decreased productivity, and higher sick-leave (Rosekind & Gregory, 2010). An effective treatment for insomnia is (computerized) cognitive behavioural therapy ((C)CBT-I) (Irwin, Cole, & Nicassio, 2006; Lancee, van den Bout, van Straten, & Spoormaker, 2012; Morin et al., 2006). It has been shown that CBT-I has better long-term effects and is better accepted by users than pharmacology (Riemann & Perlis, 2009). However, an average adherence rate of 50%, and thereby effectiveness, can still be improved (Horsch, Lancee, Beun, Neerincx, & Brinkman, 2015). This paper focuses on a tool that facilitates motivation improvement before a person starts insomnia treatment as a way to improve adherence.

Several ideas have been raised to improve adherence to behaviour change interventions. For example, people can be persuaded by all sorts of mechanisms (Beun, 2012; Fogg, 2003; Horsch, Brinkman, van Eijk, & Neerincx, 2012; Michie et al., 2013). In the health domain, the Transtheoretical Model (TTM) of behavioural change has a leading role. The TTM describes six stages through which a person progresses when changing behaviour (Prochaska & Velicer, 1997). In order to progress through the stages, first the advantages have to become more apparent to a person, then the disadvantages of change have to be perceived as less important. The balance has to tip towards the benefits before someone is ready for action (Prochaska & Velicer, 1997) (See Box 1 for an overview of all the stages).

Most interventions are tuned to people in preparation and action stage, i.e. people who are ready to change. Several studies (Noar, Benac, & Harris, 2007; Prochaska & Velicer, 1997) suggest that a person who is in the readiness-to-change stage that fits the intervention, has a higher chance of successfully completing the treatment, than people who are in another readiness-to-change stage. In other words, it is of eminent importance that people are in the preparation stage before starting an intervention.

In an assessment of current smokers it appeared that only 20% of respondents were in the preparation stage, whereas 40% of the respondents were in the precontemplation stage, and another 40% in the contemplation stage (Prochaska & Velicer, 1997). Likewise, Laforge and colleagues (1999) studied the distributions in low fat dieting, exercising, reducing stress, losing weight, and smoking in five on-going projects in Australia and the United States and found the same distribution as Prochaska and Velicer (1997). The distributions showed that the majority of people is in the maintenance stage, meaning that they already completed the behaviour change. However, from the people who had not changed yet, most people were in the precontemplation or contemplation stage.

Precontemplation	People have little intention to change and can be unaware that their behaviour is problematic. Often they are uninformed, under-informed, or demoralized.
Contemplation	People are considering change, since they are aware of the advantages of changing. Still the benefits are equally important as the disadvantages.
Preparation	People are ready to change, they have a plan of action.
Action	People are changing their behaviour.
Maintenance	People have sustained their change, they try to prevent relapse.
Termination	People are 100% confident that they can sustain their change and are never tempted to relapse anymore.

Box 1. Stages of the Transtheoretical Model

A smaller portion of people is in the preparation or action stage. For the insomnia population the same stage-distribution could probably be found as in the observed health domains. This suggests that approximately 80% of the people is not ready for action and therefore less susceptible for insomnia intervention.

Motivational interviewing (MI) is a coaching strategy that shares some fundamental ideas with the TTM (Resnicow et al., 2002). For example, in MI ambivalence about costs and benefits of changing is a central concept, which is in accordance with a shift from the precontemplation to contemplation stage. MI might be a way to align the TTM stage with a behavioural intervention. Especially people who are not ready to change can benefit from MI (Resnicow et al., 2002). In MI the ambivalence between costs and benefits of changing is a central concept, which is in accordance with the shift from the precontemplation - to contemplation - to preparation - to action stage. The effect of MI however varies. Hettema and colleagues conducted a meta-analysis including 72 studies that investigated the effect of MI on targeted health behaviour (2005). The meta-analysis showed a wide variety in effect sizes (d = -.19 to 3.25), and a meaningful average effect size could hardly be specified, even when categorized by domain (e.g., alcohol abuse, HIV studies). Still, another meta-analysis (Rubak, Sandbæk, Lauritzen, & Christensen, 2005) showed that MI sessions outperform traditional advice giving in various domains. Approximately 75% of the included studies did obtain an effect of MI on the targeted health behaviour, which Rubak and colleagues mainly attribute to improved adherence. Besides, no adverse effects of MI became apparent.

Several computerized MI interventions have been developed and studied. For example, physical activity was stimulated by different computerized MI interventions (Bickmore, Schulman, & Sidner, 2013; Blanson Henkemans et al., 2009; Di Noia, Contento, & Prochaska, 2008; Friederichs, Bolman, Oenema, Verboon, & Lechner, 2016). These studies, however, did not focus on stage changes in the TTM. To our best knowledge,

only one other study that included a computerized MI intervention measured TTM stage change (Di Noia et al., 2008). They studied fruit and vegetable intake, and found greater proportions of participants moving to a higher TTM stage in their computerized MI condition than in their control group who received a regular program.

These aforementioned computerized MI interventions have different designs and implementations. For example, Blanson Henkemans and colleagues (Bickmore et al., 2013; Blanson Henkemans et al., 2009) developed a website with a virtual character, Bickmore and colleagues (Bickmore et al., 2013) built a stand-alone virtual character, and Di Noia and colleagues (Di Noia et al., 2008) developed a CD-ROM intervention. In general, there is a trend of adding virtual characters to a wide range of applications, such as interactive learning environments (e.g., Johnson, Rickel, & Lester, 2000; Kenny et al., 2007), e-Commerce (e.g., Bauer & Neumann, 2005; Holzwarth, Janiszewski, & Neumann, 2006), and digital entertainment (e.g., Shapiro et al., 2013). This trend can also be observed in health interventions (e.g., Aziz, Bickmore, Vardoulakis, Shanahan, & Paasche-Orlow, 2012; Bickmore et al., 2013; Blanson Henkemans et al., 2009; Marsella, Johnson, & LaBore, 2003). Replacing text by a conversational agent can increase engagement (Yee, Bailenson, & Rickertsen, 2007) and thereby improve adherence (Donkin & Glozier, 2012). Substituting the text with an avatar in computerized MI interventions gave mixed results (Friederichs, Bolman, Oenema, Guyaux, & Lechner, 2014; Lisetti, Amini, Yasavur, & Rishe, 2013). In general, appreciation of the intervention increases, but the effects on the outcome measures are unclear.

Our idea was that a self-help intervention based on MI could help users progress through the different stages into the preparation stage, which eventually lead to higher success and adherence rates of CBT-I. So we developed a pre-treatment-motivation module (PMM) based on MI. Figure 1 depicts the thought process a person may go through regarding the PMM, which can be mapped to the TTM stages. To be effective a self-help intervention should address this process. A first step is the personal question whether a person has a sleep problem (precontemplation). People can be convinced that they do, or do not have a problem (step 1). However, they can also be ambivalent about it. People who believe they do not have a problem, will not be motivated to solve their problem and will exit the process remaining in the precontemplation stage. The second step is the question if a person wants to do something about the problem (contemplation). They either want to do something about it, can be ambivalent if they want to do something about it, or do not want to do something about it. For the individuals for whom step 1 led to ambivalence, step 2 is likely to lead to either ambivalence or reluctance to do anything about it. The third step is using the Pre-treatment Motivation Module (PMM) (preparation/action). If people in step 2 want to do something about their sleep problem or are ambivalent about it, they can use the PMM (step 3) and come to a conclusion about tackling their sleep problem (step 4).

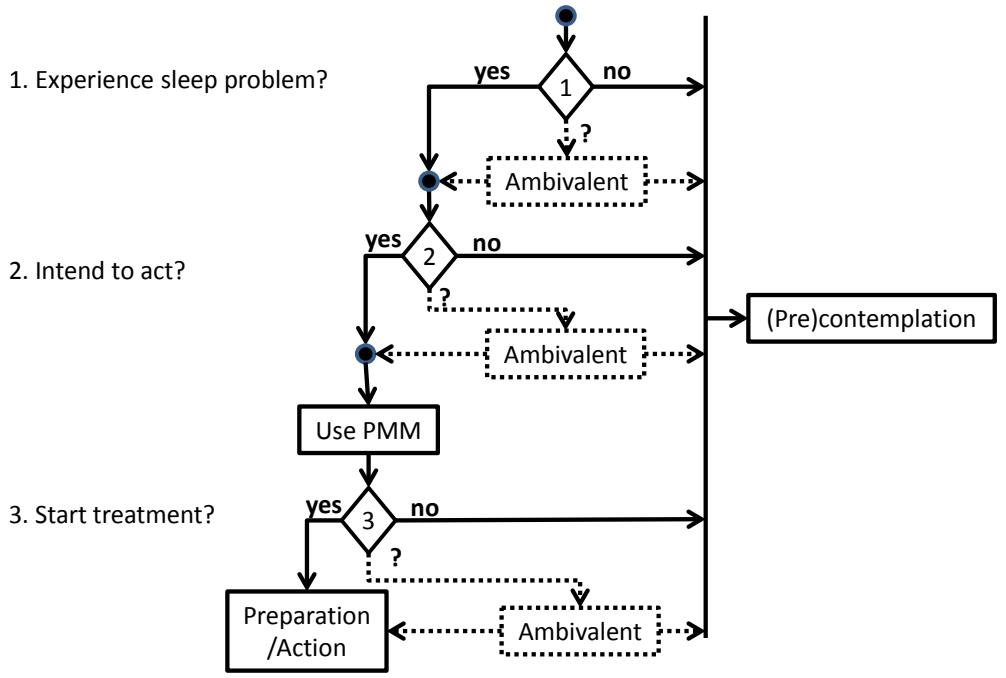


Figure 1. Thought process

3.2. The experiment

The primary goal of the experiment was to find out whether people can progress from one to another TTM-stage by a self-help intervention that was based on the ideas of MI. Therefore a stage-matched intervention was developed, called the PMM. A secondary goal was to explore if an interactive coach was better able to help people in that process than a paper and pencil workbook. For this a mixed experiment was set up with within-subject pre-post measures, and between-subject paper vs interactive tool type.

3.2.1. Intervention

The PMM was a self-help workbook concerning sleep problems that helped participants structuring their thoughts and reflect on them. The PMM consisted of seven chapters and 34 MI exercises. Exercises were for example, reading and learning about sleep and sleep problems, thinking about what is important in one's life (values), and writing an action plan. Participants could decide themselves which exercises they wanted to do. The

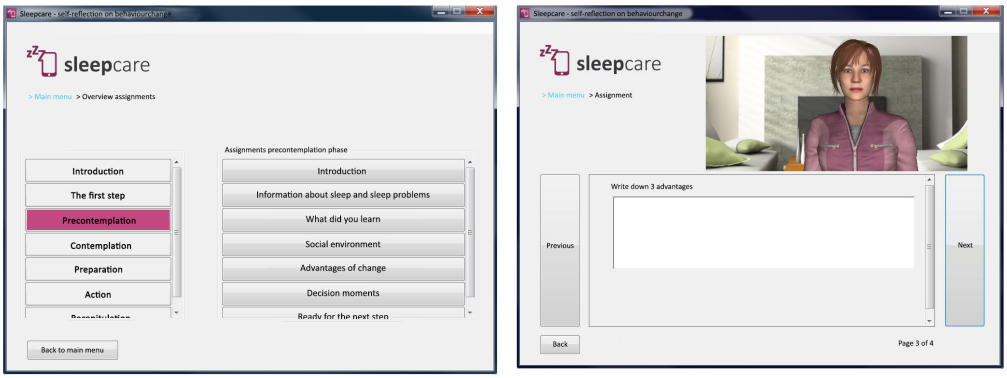


Figure 2. Screenshot of the computer version of the PMM (translated from Dutch)

chapters represented the TTM stages precontemplation, contemplation, preparation, and action. There were also an introduction, assessment, and recap chapter. At the end of every chapter participants could fill in a questionnaire which referred them to the chapter that matched their current TTM stage.

Two versions of the PMM were developed: a paper version and a computer version including an avatar coach. The paper workbook was an A4 bound booklet. The computer version included a female avatar as depicted in Figure 2. She vocally explained the exercises and gave feedback after the participants typed in their answer. The feedback was based on the number of words used by the participants and the time the participants spent on their answer. If participants spent enough time or wrote an extensive answer the avatar gave a compliment, encouraged them to keep on going, and/or showed her appreciation. If a participant's answer was insufficient (spent too little time, or wrote too few words) the avatar either a) repeated the goal of the exercise, b) repeated the explanation of the exercise in different words, c) asked the participant a clarifying question or gave extra information, d) emphasized the importance of the exercise, or e) expressed its faith in the participant. Feedback was always stated in a positive way. A print screen of the computer version can be seen in Figure 2.

3.2.2. Procedure

Participants for the experiment were recruited via social media, online advertisements, the website 'www.ikgalekkerslapen.nl', and in university lectures. After giving online consent, participants completed the first online questionnaire that measured TTM-stage (URICA) (DiClemente & Hughes, 1990), Insomnia Severity Index (ISI) (Bastien, Vallières, & Morin, 2001), and general questions about preferences and self-efficacy. Next, the participants were randomly assigned to the paper or digital PMM, which was mailed to their home. The assignment was to work at least twice with the PMM within a period of

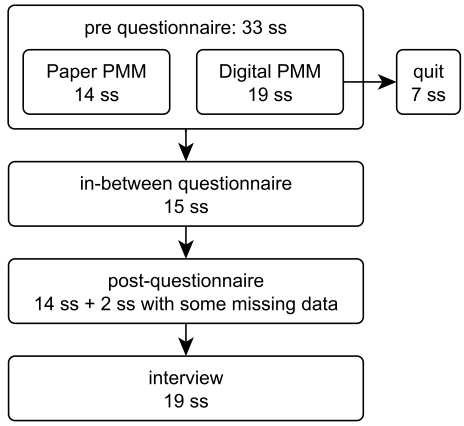


Figure 3. Flowchart of participants

four weeks. After two weeks they received an e-mail with an in-between questionnaire regarding the acceptance of the PMM based on the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh, Morris, Davis, & Davis, 2003). This e-mail also served as an implicit reminder. Again two weeks later participants received the post-questionnaire consisting of the URICA, ISI, motivation questionnaire (SIMS) (Guay, Vallerand, & Blanchard, 2000), specific adherence questions, and open questions about the effect of the workbook and the perceived effort of the participants. In addition, participants were contacted for a telephone interview. The study was approved by the internal Ethical Review Board of the Delft University of Technology.

3.2.3. Participants

In total 33 people participated in the experiment (Figure 3). 14 participants received the paper workbook, and 19 people received the digital workbook on a USB-stick. Seven participants e-mailed that they wanted to quit, all of them had the digital workbook. Reasons for quitting were for example, illness, accident, or other solution found. 15 participants filled in the in-between (UTAUT) questionnaire; 16 participants filled in the post questionnaire, with two of them having some missing data. So, around 50% of the participants filled in the questionnaires. 19 people participated in the interview.

3.2.4. Experimental findings

Two of the 15 people from whom we had both pre- and post- URICA questionnaire progressed from the precontemplation stage to the contemplation stage, four participants showed the reverse transition from contemplation, to the precontemplation stage. The other nine people remained in the same stage. For eight people that meant staying in the contemplation stage, and for one person it meant staying in the precontemplation

Table 1. Means and Standard Deviations of the different measures

	Pre	Post
Precontemplation	*n* = 3	*n* = 4
Contemplation	*n* = 13	*n* = 11
TTM-score	*M* = 68.47, *SD* = 10.83, *n* = 15	*M* = 60.43, *SD* = 10.61, *n* = 14
ISI	*M* = 16.57, *SD* = 5.23, *n* = 14	*M* = 12.36, *SD* = 5.03, *n* = 14
Easiness-To-Initiate		*M* = 2.67, *SD* = 1.41, *n* = 15
Intrinsic Motivation		*M* = 3.61, *SD* = 1.31, *n* = 14
Identified Regulation		*M* = 4.77, *SD* = 1.26, *n* = 14
External Regulation		*M* = 2.98, *SD* = 1.33, *n* = 14
Amotivation		*M* = 2.59, *SD* = 1.13, *n* = 14

stage. To an open question that asked what kind of follow-up actions participants had taken, five of the 15 people answered that they took some sort of action as a result of the workbook, e.g., searching for more information, went to the general practitioner for sleep medication, or seeing a psychologist.

There was a significant difference in sleep severity (ISI) before (*M* = 16.57, *SD* = 5.23, *n* = 14) and after intervention (*M* = 12.36, *SD* = 5.03, *t*(13) = 3.71, *p* = .003, *d* = .82) (Table 1). On average people suffer less from insomnia after using the PMM. Three of those people achieved a clinically meaningful difference, which is set on an 8-point reduction (Morin, Belleville, Bélanger, & Ivers, 2011).

From 23 participants we had the post-questionnaire data and/or interview data. This combined data suggests that half of the participants (11 of the 23) used the workbook at least twice, as was asked of them. The other half of the participants only used it once or never. Participants found it difficult to start using the workbook. Four 7-point statements using an equal interval were used to measure the easiness to initiate working with the workbook. The average score was low (*M* = 2.7, *SD* = 1.4, *n* = 15). Furthermore, the data suggests that participants were mediocre motivated to use the PMM. Motivation was measured from 1 to 7 with higher scores indicating higher motivation. The scale distinguishes Intrinsic Motivation (IM) (*M* = 3.6, *SD* = 1.3, *n* = 14), Identified Regulation (IR) (*M* = 4.8, *SD* = 1.3, *n* = 14), External Regulation (ER) (*M* = 3.0, *SD* = 1.3, *n* = 14), Amotivation (AM) (*M* = 2.6, *SD* = 1.1, *n* = 14). Because a limited number of participants (*n* = 15) completed the acceptance (UTAUT) questionnaire, no analysis was done on that dataset. Furthermore, we also refrain from the second research aim, investigating the difference between paper-computer, because of a lack of data.

Ambivalence about sleep

"I would like to sleep better, but I am not sure what my exact problem is." (ss 15, Female, 31 years)

" I lie awake now and then, but I am not that worried about it." (ss 18, Female, 51 years)

"That I don't sleep is caused by external things, however, if I would sleep deeper, I probably won't be awakened by it." (ss 8, Female, 49 years)

Ambivalence about workbook usage

"I understand the structure and reflection part, but I just want to start doing sleep exercises" (ss 20, Female, 30 years)

"I think it is a pity, especially for me personally, that I did not spend more time on it." (ss 23, Male, 48 years)

"I had good intentions, but I just failed" (ss 10, Female, 48 years)

"There were so many exercises, and although it states that you can choose what to do, I actually want to do it all." (ss 20, Female, 30 years)

"At the moments on which I had an open mind, it was easy to start working with it." (ss 17, Female, 25 years)

Box 2. Example quotes regarding ambivalence

3.2.4.1. Interviews

The interviews were audio recorded and transcribed by the first author (CH). The transcripts were read and statements regarding ambivalence and change talk were singled out. Two different kinds of ambivalence surfaced. Users were not only ambivalent about their sleep and sleep problem, as expected, but also about the PMM, which was not foreseen (Box 2). Users' ambivalence about the pre-treatment motivation module could possibly explain the scant usage of the PMM.

3.3. Interpretation of the findings and possible explanations

The PMM was a stage-matched intervention aimed at dealing with ambivalence. The results, however, showed not only ambivalence about sleep, but also about the PMM itself. This result suggests an extra step in a model that describes the thought paths people could follow (Figure 1).

Figure 4 depicts two models. On the left side the naïve model as described before, and on the right side the model derived from the data. In the middle the consideration steps are depicted. The naive model assumes that people who want to do something about their sleep problem (step 2) will use the PMM (step 3), and people who do not want to do something about the problem (step 2), will not use the PMM (step 3). People who are ambivalent about doing something about their sleep problem (step 2) could either use the PMM or not (step 3). The extended model, however, suggests there exists an

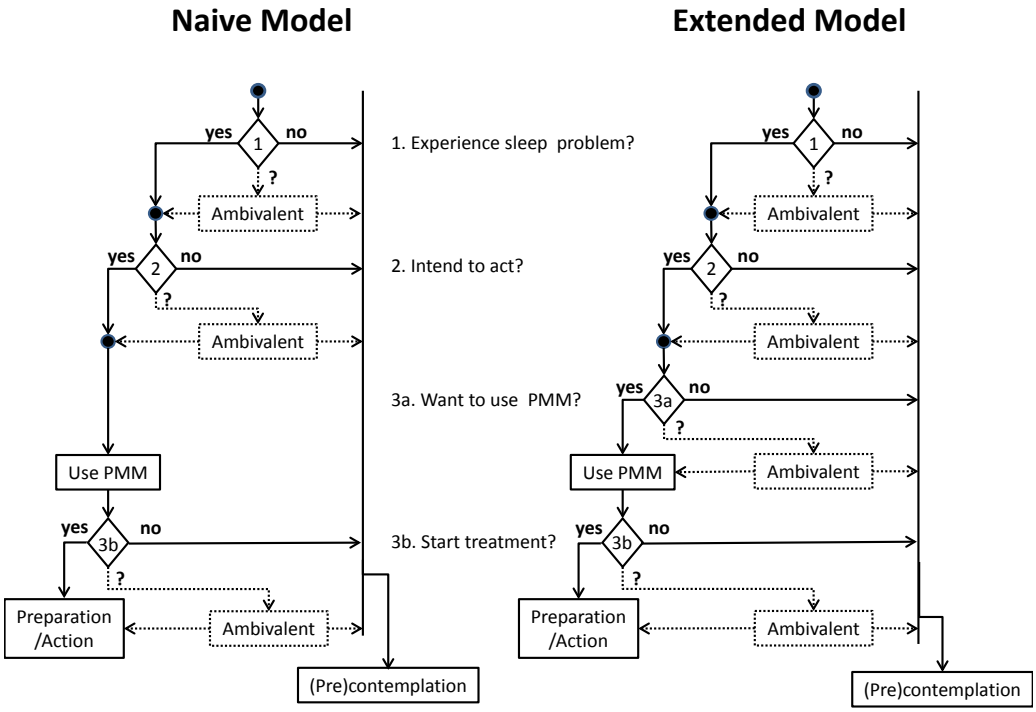


Figure 4. The naïve model and the data-driven model describing the thought process regarding the PMM

additional step (3a) between these two steps. People who want to do something about their sleep problem could still be ambivalent or aversive towards the PMM resulting in non-usage. This extra step generates additional paths that could partly explain non usage of the PMM.

Although in retrospect the extra step seems somewhat trivial, ambivalence about tools or interventions has not been given much attention. In the TTM and MI the focus lies on ambivalence about health problems and how to advance through that process, whereas other behaviour change theories explain if and when target behaviour occurs. To our knowledge current literature does not explicitly describe behaviour change theories that focus on tool ambivalence. For example, take well-known behaviour (change) theories like the theory of planned behaviour (Ajzen, 1991), theory of reasoned action (Ajzen & Fishbein, 1988), social-cognitive theory (Bandura, 1977), cognitive dissonance theory (Festinger, 1957), the COM-B model (Michie, van Stralen, & West, 2011; Ripple, 1955), or the BMAT model (Fogg, 2009). These theories explain how and when behaviour occurs, but none of them include the concept of tool ambivalence.

Tool ambivalence may be seen as a lack of motivation to use that tool, in the context of those known behaviour theories. Miller and Rollnick (2003) object against this way of thinking, since they consider ambivalence to be a natural process when changing behaviour. However, considering tool ambivalence as a lack of motivation might help to avoid the downward-spiral that arises when only thinking in terms of ambivalence. When non-adherence is explained in terms of ambivalence a downward-spiral arises; being ambivalent about a health problem – being ambivalent about an intervention – being ambivalent about a motivation-enhancing tool (e.g., the PMM) intended to decrease ambivalence about the intervention – being ambivalent about an information brochure about the tool – etc.

A similar process has emerged in procrastination research. Recently Andreou introduced the term second-order procrastination, defined as '*procrastinating in dealing with a (first-order) procrastination problem*' (Andreou, 2007, p. 1). Ambivalence can be considered analogously to first- and second-order procrastination, in which ambivalence about a health problem matches first-order procrastination, and ambivalence about the tool compares to second-order procrastination. A simple explanation of second-order procrastination offered by Andreou is that the costs are too high for making a plan to solve the (first-order) procrastination. Similarly, one could argue that if people are ambivalent about a tool or intervention the costs are too high and the expected benefits are too low.

Decision theory (Schlaifer & Raiffa, 1961) framed this problem as the expected value of sample information (EVSI). Additional information from a sample allows people to make a more informed decision. Samples could be the weather forecast, advice of a consultant, or the results of an experiment. In our case the sample is the PMM, which aims to reduce ambivalence and increase knowledge about CBT-I. EVSI estimates the extra value of the sample data before actually having that sample data, so before using the PMM. If the costs of gathering the sample information exceeds the EVSI, rational people will decide not to invest in a sample. In our case, if the costs of using the PMM is higher than the expected value of the knowledge gained by the PMM, EVSI predicts that people will not use the PMM.

Learning theory would label this problem as the production paradox (Carroll & Rosson, 1987). The production paradox's starting point is that people's main goal is to get something done (the active user). Since learning something often does not directly lead to the desired output, people are not inclined to put the effort into thorough learning. This may lead, for example, to situations in which people stick to already known procedures instead of learning new, more effective procedures. Translated to people with a sleep problem and their use of the PMM, the production paradox implies that people are not inclined to use the PMM and learn about themselves and their sleep problem, but just want to get to a quick fix for their sleep problem.

3.4. Discussion

Two versions of a pre-treatment motivation module were developed to help people progress through the stages of change. The ultimate aim of the PMM was increasing adherence to CBT-I. Paradoxically, this study showed difficulties with adherence to the PMM. From the data an extended model was derived that showed people are not only ambivalent about their possible problem, but also about the PMM itself.

Although adherence to the PMM did not meet our expectations, a PMM might still have an effect on a subsequent CBT-I intervention. Marino and colleagues (2010) studied the effect of a single-session MI intervention delivered by telephone on the attendance to group CBT-I treatment. Although the study sample was too small to draw any significant conclusions, they observed that five out of the ten participants who received a MI-based telephone call attended the group session. Whereas, only one of the eight people from the control group was present at the group session (Marino, 2010). So, it might be that using the PMM only once will increase adherence to a CBT-I intervention. Which leads to the first limitation of this study.

The first limitation of this study was the lack of treatment after the PMM. If CBT-I was offered after participants used the PMM, the ultimate goal of increasing adherence in CBT-I could have been investigated. The primary goal of this study, however, was to explore if people could progress through the TTM stages, and the secondary goal was to find differences between the paper and interactive version of the PMM. The ultimate goal of improving adherence in CBT-I by the PMM was part of future work. Another limitation of this study was the lack of control. The PMM was sent to the participants' homes and they were responsible to work with it themselves. This freedom corresponds to the freedom people have when helping themselves in daily life, which increases the ecological validity of this experiment. Besides, the main contribution of this study is the finding that users are also ambivalent about tools. This finding arose from the experimental set-up in the field. If, participants would have come to the lab, ambivalence about the PMM would not have been discovered.

Possible solutions to non-usage can be found in decision theories, learning theories, behaviour theories, or motivation theories. First, the EVSI suggests that the costs of using the PMM should be lowered, or the perceived benefits should be increased. The production paradox suggests more or less the same solution. Carroll and Rosson (1987) propose that learning itself should be made less demanding (lowered costs), and intrinsic rewards should be made more salient (increase perceived benefits). Furthermore, they suggest that people are goal-orientated and that this could be used as a drive for learning, or in this case for using the PMM. To cater to people's goal orientation, the goal of each exercise was explained at the beginning of the exercise. Furthermore, one of the design rationales of the PMM was that users could freely browse through the exercises, which also accommodates people's goal orientation. Other ideas to use people's goal orientation as a drive could reinforce usage. Another strategy would be to focus on the problem that people simply forget to use the intervention, following Fogg's

behaviour wizard the intervention can be seen as span behaviour, which is characterized by usage that is limited to specific duration (Fogg & Hreha, 2010). Fogg and Hreha (2010) suggest that span behaviours benefit from regular triggers, which were lacking in this experiment. Yet another way to handle tool ambivalence and non-adherence of the PMM is increasing extrinsic motivation. For long-term behaviour change, such as following CBT-I, intrinsic motivation is preferred over extrinsic motivation (Deci, Koestner, & Ryan, 1999). The PMM exactly targets improving intrinsic motivation by MI. However, for the PMM itself it is not necessary to increase intrinsic motivation and extrinsic motivation can be adequate. Extrinsic motivation could for example be improved by persuasion or gamification.

3.5. Conclusion

The goal of this paper was to study the effects of a pre-treatment motivation module on the Trans Theoretical Model stage of people. An additional goal was to explore the difference between a paper and an electronic version of the Pre-treatment Motivation Module. The study was a field experiment in which participants were treated as if they received self-help.

The results showed that most participants were in the contemplation stage and remained in that stage over the course of 4-week experiment. Moreover, the findings showed that approximately half of the participants did not use the Pre-treatment Motivation Module as much as was required. This could possibly be explained by tool ambivalence. The study's main scientific contribution is the new established concept of tool ambivalence.

Considering ambivalence as the root of non-usage leads to a downward spiral, which is hard to break. Therefore, we also framed the observed phenomenon as a decision problem. By framing it in this way, non-usage could be explained by the expected value of sample information and the cost to get this information. Possible ways to solve non-usage of the Pre-treatment Motivation Module are increasing extrinsic motivation or adding reminders.

3.6. References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211.

Ajzen, I., & Fishbein, M. (1988). Theory of reasoned action-Theory of planned behavior. *University of South Florida*.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed ed.). Arlington, VA: American Psychiatric Publishing.

Andreou, C. (2007). Environmental preservation and second-order procrastination. *Philosophy & Public Affairs, 35*(3), 233-248.

Aziz, M., Bickmore, T. W., Vardoulakis, L. P., Shanahan, C., & Paasche-Orlow, M. (2012). *Using computer agents to improve emotional wellbeing of hospital patients*. Paper presented at the CHI'12, Austin, Texas, USA.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review, 84*(2), 191.

Bastien, C. H., Vallières, A., & Morin, C. M. (2001). Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep medicine, 2*(4), 297-307.

Bauer, H. H., & Neumann, M. M. (2005). *Investigating the effects of avatars as virtual representatives in electronic commerce.* Paper presented at the Proceedings of the Australian-New Zealand Marketing Academy Conference (ANZMAC).

Beun, R. J. (2012). Persuasive strategies in mobile insomnia therapy: alignment, adaptation, and motivational support. *Personal and Ubiquitous Computing*, 1-9. doi: 10.1007/s00779-012-0586-2

Bickmore, T. W., Schulman, D., & Sidner, C. (2013). Automated interventions for multiple health behaviors using conversational agents. *Patient Educ Couns, 92*(2), 142-148.

Blanson Henkemans, O. A., van der Boog, P. J. M., Lindenberg, J., van der Mast, C. A. P. G., Neerincx, M. A., & Zwetsloot-Schonk, B. J. H. M. (2009). An online lifestyle diary with a persuasive computer assistant providing feedback on self-management. *Technology and Health Care, 17*(3), 253-267.

Carroll, J. M., & Rosson, M. B. (1987). *Paradox of the active user*: The MIT Press.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol Bull, 125*(6), 627.

Di Noia, J., Contento, I. R., & Prochaska, J. O. (2008). Computer-mediated intervention tailored on transtheoretical model stages and processes of change increases fruit and vegetable consumption among urban African-American adolescents. *American Journal of Health Promotion, 22*(5), 336-341.

DiClemente, C. C., & Hughes, S. O. (1990). Stages of change profiles in outpatient alcoholism treatment. *Journal of substance abuse, 2*(2), 217-235.

Donkin, L., & Glozier, N. (2012). Motivators and Motivations to Persist With Online Psychological Interventions: A Qualitative Study of Treatment Completers. *J Med Internet Res, 14*(3).

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Fogg, B. J. (2003). *Persuasive technology: using computers to change what we think and do*. San Francisco: Morgan Kaufmann Publishers.

Fogg, B. J. (2009). *A behavior model for persuasive design.* Paper presented at the Persuasive 2009, Claremont, California, USA.

Fogg, B. J., & Hreha, J. (2010). *Behavior wizard: a method for matching target behaviors with solutions*. Paper presented at the Persuasive Technology.

Friederichs, S., Bolman, C., Oenema, A., Guyaux, J., & Lechner, L. (2014). Motivational interviewing in a web-based physical activity intervention with an avatar: Randomized controlled trial. *J Med Internet Res, 16*(2).

Friederichs, S., Bolman, C., Oenema, A., Verboon, P., & Lechner, L. (2016). Exploring the working mechanisms of a web-based physical activity intervention, based on self-determination theory and motivational interviewing. *Internet Interventions, 3*, 8-17.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). *Motivation and emotion, 24*(3), 175-213.

Hettema, J., Steele, J., & Miller, W. R. (2005). Motivational interviewing. *Annu. Rev. Clin. Psychol., 1*, 91-111.

Holzwarth, M., Janiszewski, C., & Neumann, M. M. (2006). The influence of avatars on online consumer shopping behavior. *Journal of marketing, 70*(4), 19-36.

Horsch, C., Brinkman, W. P., van Eijk, R., & Neerincx, M. (2012). *Towards the usage of persuasive strategies in a virtual sleep coach*. Paper presented at the UK HCI 2012 Workshop on People, Computers & Psychotherapy, Birmingham.

Horsch, C. H. G., Lancee, J., Beun, R. J., Neerincx, M. A., & Brinkman, W. P. (2015). Adherence to Technology-mediated Insomnia Treatment: A meta-analysis, interviews with users, and focus groups with users and experts. *Journal of Medical Internet Research, 17*(9), e214. doi: 10.2196/jmir.4115

Irwin, M. R., Cole, J. C., & Nicassio, P. M. (2006). Comparative meta-analysis of behavioral interventions for insomnia and their efficacy in middle-aged adults and in older adults 55+ years of age. *Health Psychol, 25*(1), 3-14. doi: 10.1037/0278-6133.25.1.3

Johnson, W. L., Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education, 11*(1), 47-78.

Kenny, P., Hartholt, A., Gratch, J., Swartout, W., Traum, D., Marsella, S., & Piepol, D. (2007). *Building interactive virtual humans for training environments.* Paper presented at the Proceedings of I/ITSEC.

Laforge, R. G., Velicer, W. F., Richmond, R. L., & Owen, N. (1999). Stage distributions for five health behaviors in the United States and Australia. *Preventive medicine, 28*(1), 61-74.

Lancee, J., van den Bout, J., van Straten, A., & Spoormaker, V. I. (2012). Internet-delivered or mailed self-help treatment for insomnia? A randomized waiting-list controlled trial. *Behaviour Research and Therapy, 50*(1), 22-29. doi: 10.1016/j.brat.2011.09.012

Lisetti, C., Amini, R., Yasavur, U., & Rishe, N. (2013). I can help you change! an empathic virtual agent delivers behavior change health interventions. *ACM Transactions on Management Information Systems (TMIS), 4*(4), 19.

Marino, C. (2010). *Motivational interviewing in insomnia treatment: a randomized control pilot study*: PALO ALTO UNIVERSITY.

Marsella, S., Johnson, W. L., & LaBore, C. (2003). *Interactive pedagogical drama for health interventions.* Paper presented at the 11th International Conference on Artificial Intelligence in Education, Sydney, Australia.

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., . . . Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of behavioral medicine, 46*(1), 81-95.

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation Science, 6*(1), 42.

Miller, W., & Rollnick, S. (2003). Motivational interviewing: Preparing people for change. *Journal for Healthcare Quality, 25*(3), 46.

Morin, C. M., Belleville, G., Bélanger, L., & Ivers, H. (2011). The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep, 34*(5), 601.

Morin, C. M., Bootzin, R. R., Buysse, D. J., Edinger, J. D., Espie, C. A., & Lichstein, K. L. (2006). Psychological and behavioral treatment of insomnia: update of the recent evidence (1998-2004). *SLEEP-NEW YORK THEN WESTCHESTER-, 29*(11), 1398.

Noar, S. M., Benac, C. N., & Harris, M. S. (2007). Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol Bull, 133*(4), 673-693. doi: 10.1037/0033-2909.133.4.673

Ohayon, M. M. (2002). Epidemiology of insomnia: what we know and what we still need to learn. *Sleep medicine reviews, 6*(2), 97-111.

Prochaska, J. O., & Velicer, W. F. (1997). The transtheoretical model of health behavior change. *American journal of health promotion, 12*(1), 38-48.

Resnicow, K., DiIorio, C., Soet, J. E., Borrelli, B., Hecht, J., & Ernst, D. (2002). Motivational interviewing in health promotion: It sounds like something is changing. *Health Psychology, 21*(5), 444-451.

Riemann, D., & Perlis, M. L. (2009). The treatments of chronic insomnia: a review of benzodiazepine receptor agonists and psychological and behavioral therapies. *Sleep Medicine Reviews, 13*(3), 205-214.

Ripple, L. (1955). Motivation, capacity, and opportunity as related to the use of casework service: Theoretical base and plan of study. *Social Service Review, 29*(2), 172-193.

Rosekind, M. R., & Gregory, K. B. (2010). Insomnia risks and costs: health, safety, and quality of life. *Am J Manag Care, 16*(8), 617-626.

Rubak, S., Sandbæk, A., Lauritzen, T., & Christensen, B. (2005). Motivational interviewing: a systematic review and meta-analysis. *The British journal of general practice, 55*(513), 305.

Schlaifer, R., & Raiffa, H. (1961). Applied statistical decision theory.

Shapiro, D. G., McCoy, J., Grow, A., Samuel, B., Stern, A., Swanson, R., . . . Mateas, M. (2013). *Creating Playable Social Experiences through Whole-Body Interaction with Virtual Characters.* Paper presented at the AIIDE.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

Yee, N., Bailenson, J. N., & Rickertsen, K. (2007). *A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces.* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.

4. Reminders make people adhere better to a self-help sleep intervention

Abstract

Objective - The experiment presented in this paper investigated the effects of different kind of reminders on adherence to automated parts of a cognitive behavioural therapy for insomnia (CBT-I) delivered via a mobile device.

Background - Previous studies report that computerized health interventions can be effective. However, treatment adherence is still an issue. Reminders are a simple technique that could improve adherence. A minimal intervention prototype in the realm of sleep treatment was developed to test the effects of reminders on adherence. Two prominent ways to determine the reminder-time are: a) ask users when they want to be reminded, and b) let an algorithm decide when to remind users.

Method - The prototype consisted of a sleep diary, a relaxation exercise and reminders. A within subject design was used in which the effect of reminders and two underlying principles were tested by 45 participants that all received the following three different conditions (in random order): a) event-based reminders b) time-based reminders c) no reminders.

Results - Both types of reminders improved adherence compared to no reminders. No differences were found between the two types of reminders. Opportunity and self-empowerment could partly mediate adherence to filling out the sleep diary, but not to the number of relaxation exercises conducted.

Contribution - Although the study focussed on CBT-I, we expect that designers of other computerized health interventions benefit from the tested opportunity and self-empowerment principles for reminders to improve adherence, as well.

Chapter accepted as:

Horsch, C. H. G., Spruit, A. G. L., Lancee, J., Eijk, R. M. v., Beun, R. J., Neerincx, M. A., & Brinkman, W. P. (accepted). Reminders make people adhere better to a self-help sleep intervention. *Health and Technology*.

4.1. Introduction

Everyone forgets to do something now and then. The consequences of forgetting to do something depend on what was forgotten. Some of these memory failures can be fatal. Fogetting to buy milk is less problematic than forgetting to take a sleeping baby out of a soon to be hot car, or medical errors during surgery. The intention to do something in the future is formed in prospective memory (McDaniel & Einstein, 2007). Everyone suffers from prospective memory failures. In fact, Kliegel and Martin (Kliegel & Martin, 2003) state that 50-80% of everyday life forgetting is due to these prospective memory failures. Another study in the health domain found that most of the preventable mistakes were prospective memory failures (Rothschild et al., 2005). Not only medical professionals suffer from prospective memory failures, also patients suffer from it. For example, people forget to take their pills. Forgetfulness, or prospective memory failure, is one of the main reasons (Jones, Darroch, & Henshaw, 2002; Smith & Oakley, 2005; Unni & Farris, 2011).

Several models attempt to explain how prospective memory works e.g., the preparatory attentional and memory theory, the reflexive-associative theory, and the multi-process model (McDaniel & Einstein, 2007). The latter two theories include cues. The idea is that an intended action is associated with a cue. When that cue occurs the intended action is remembered automatically. Reminders provided for example by a smartphone can serve as these cues and might play an important role in performing targeted behaviour.

Reminders have been used in various domains in different forms for a long time. Ranging from tying a string around your finger, self-written notes, to reminders set on PDAs, watches, and smartphones. Fogg describes three types of triggers in his behavioural model (Fogg, 2009). He distinguishes sparks, facilitators, and signals. A spark is a cue that enhances motivation. There are three core motivators that sparks can use: pleasure-pain, hope-fear, and social acceptance-rejection. A facilitator is a cue that makes it easier to exhibit a certain type of behaviour (enhances ability). A signal is a simple reminder used in cases where both motivation and ability are high. Another distinction in reminders can be made based on the trigger method utilized. Various trigger methods are time-based, event-based, and location-based. An example of a time-based action is taking cookies out of the oven in 20 minutes, an example of an event-based action is bringing up an issue during the next meeting, and a location-based action is throwing a letter in a mailbox when passing by. Prospective memory also makes use of these type of triggers, and research has shown that people perform better at event-based intentions than at time-based intentions (Dismukes, 2012). Especially if target behaviour has to be performed at a specific time, people could benefit from a reminder system.

Furthermore, interruptibility has been studied extensively (e.g., Ho & Intille, 2005; Hudson et al., 2003; Turner, Allen, & Whitaker, 2015). Traditionally, task complexity, task duration, and the moment of interruption has been identified as determining factors for the appropriateness of an interruption (Pejovic, Musolesi, & Mehrotra, 2015). Recently, mobile interruptibility studies shifted the focus to the moment of interruption. Mobile

studies have shown that smartphone notifications can have inappropriate timing (Fischer, Greenhalgh, & Benford, 2011; Smith, Lavygina, Ma, Russo, & Dulay, 2014). In mobile interruptibility studies context is often mentioned as the determining factor for the appropriateness of an interruption. Context, however, is a comprehensive concept that is used differently in studies. For example, Ho and Intille (2005) measure physical activity and appropriateness of interruptions, whereas Pielot and colleagues (Pielot, De Oliveira, Kwak, & Oliver, 2014) use phone usage data to infer interruptibility. Independent of the definition of context that is used, all studies acknowledge the importance of appropriate timing.

There is substantial evidence that computerized health interventions can be effective (Andersson, 2010; Kaltenthaler & Cavanagh, 2010; Wantland, Portillo, Holzemer, Slaughter, & McGhee, 2004). However, adherence remains a challenge. Compared to more traditional treatments, computerized interventions can be experienced as less binding, therefore it is easy to drop-out (Eysenbach, 2005). Since, the efficacy of treatment is partly determined by adherence (Donkin et al., 2011; Gould & Clum, 1993; Horsch, Lancee, Beun, Neerincx, & Brinkman, 2015; World Health Organization, 2003) it is crucial to optimize adherence (Edinger & Means, 2005). One of the reasons why people do not adhere to health interventions is forgetfulness (Donkin & Glozier, 2012; Horsch et al., 2015). Reminders are a simple technique that could help solving this particular problem of forgetfulness (Donkin & Glozier, 2012). For example, earlier studies in the health domain have shown that mobile text reminders increase show-up rates for gastrointestinal endoscopy (Deng et al., 2015), for breast cancer screening (Vidal et al., 2014), and sunscreen use (Armstrong et al., 2009). Another example, regarding an app with notifications, showed an increase from 12% to 63% in logging food intake on a mobile phone when reminders were given compared to the absence of reminders (Bentley & Tollmar, 2013). Moreover, a systematic review about reminders in cell phone interventions found a difference in 20 of the 25 studies between the intervention and control group (Krishna, Boren, & Balas, 2009). This indicates that reminders improve adherence and the outcome of interventions.

Previous research suggests that reminders can be effective, to our knowledge, however, barely any empirical work has been done regarding the underlying principles that explain why these reminders work. This paper discusses and tests two prominent reminder approaches: time-based reminders and reminders inspired by the interruptibility literature. The time-based reminders are self-set reminders in which the user can choose the time. The other reminders are automatic event-based reminders, inspired by the Capability-Opportunity-Motivation Behaviour (COM-B) model (Michie, van Stralen, & West, 2011). Here the system detects opportune moments and sends a reminder.

A domain that can benefit from effective reminders is mobile insomnia treatment. People who suffer from insomnia have difficulties initiating or maintaining sleep (Morin, 1993). Having insomnia leads to personal suffering, like a reduced quality of life, and societal costs, like reduced productivity (Rosekind & Gregory, 2010). Studies

estimate that about 10% of the adults suffer from insomnia (Ohayon, 2002). Cognitive-Behaviour Therapy for Insomnia (CBT-I) is the treatment of choice for this disorder and fairly standardized in protocols (Espie et al., 2007). CBT-I consists of several exercises that requires behaviour changes, however, adherence to CBT-I remains a problem (Dimatteo, Giordani, Lepper, & Croghan, 2002; Horsch et al., 2015). For instance, a daily sleep diary that helps people to become aware of their sleep behaviour and monitor progress is easily missed. Reminding people to do their exercises could be beneficial and provides opportunity to test the effect and underlying principles of reminders.

4.2. Reminder design and hypotheses

Earlier work has shown that reminders probably work, but it might depend on the domain, the patient demographics, psychosocial and behavioural characteristics, etc. (Fenerty, West, Davis, Kaplan, & Feldman, 2012; Krueger, Felkey, & Berger, 2003) Therefore, the first step is investigating if reminders in an sleep intervention domain, delivered via a smartphone are effective. So, hypothesis 1 is:

H1: Reminders increase adherence compared to no reminders. When people are reminded to do something they will do it more often, compared to situations in which they are not reminded to do it. There was no hypothesis regarding an adherence difference between the two types of reminder.

Figure 1 depicts this and the following hypotheses regarding the effect of reminders and their underlying mechanisms.

4.2.1. Self-set reminder

Having users set the reminder times themselves (instead of automatically set reminders) might be an adequate mechanism, because users know best when they have time to perform an activity. Several psychological principles support and explain why self-set reminders increase adherence. Firstly, the self-determination theory (Ryan & Deci, 2000) states that supporting autonomy, competence, and relatedness increases people's motivation and performance. By giving users the control to set the reminder times, their autonomy is supported and thereby their motivation and performance increases. Secondly, the consistency principle (Festinger, 1957) states that humans want to be consistent in their attitudes, beliefs, perceptions, and behaviours and they will change any of these when inconsistency is discovered. This suggests that, when people set the reminder times themselves, they are more likely to follow-up on those reminders, because they want to be consistent. Thirdly, predictability could be seen as a basic human motive. From an evolutionary viewpoint, higher predictability of (dangerous) events gives a better chance of survival. Campbell and Tesser (1983)

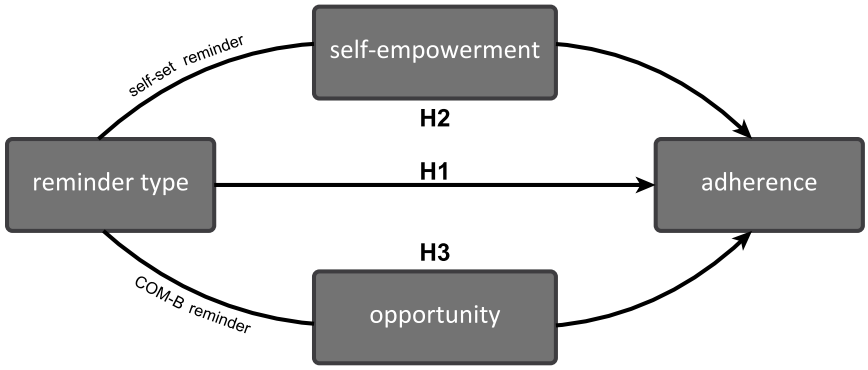


Figure 1. The three hypothesized relationship between the type of reminder, the explaining concepts and adherence.

construct the predictability motive from the human desire for certainty, the need for an understandable world, and the need to be able to predict the environment. In line with this predictability motive, reminders will be perceived more positive when arriving at predictable times, and adherence will benefit from this positive attitude.

In conclusion, self-set reminders could help to improve adherence because of three underlying principles. First, users would probably feel more in control and therefore respond more positive to the reminders. Second, people might feel committed to their self-set reminders, which would also elicit more positive responses. Third, the reminders are more predictable when the users set the times themselves, and this should also improve the response to reminders. In summary, these self-set reminders should increase the sense of self-empowerment, and therefore increase treatment adherence.

H2: If self-set reminders are given, self-empowerment mediates adherence. When people receive reminders at times they set themselves, they feel more empowered. In which self-empowerment includes, perceived control, commitment and predictability. So, it was hypothesised that self-set reminders increase perceived self-empowerment, and that perceived self-empowerment was associated with their adherence. Therefore, self-empowerment could partly explain adherence, when self-set reminders were given.

4.2.2. COM-B reminder

Reminders can also be triggered automatically. The Capability-Opportunity-Motivation-Behaviour (COM-B) model states that the possibility that people exhibit a behaviour depends on the capability of a person, the opportunity, and their motivation to exhibit that behaviour (Michie et al., 2011). Opportunity in this context means the circumstances

that allow someone to exhibit the targeted behaviour. For example, taking the stairs, instead of the elevator, is only possible when there are stairs (opportunity). Besides, it is easier to climb stairs when they are located in front of the entrance (opportune location), or when colleagues take the stairs (social opportunity). The model suggests that if people are capable and motivated to exhibit the targeted behaviour, a reminder at an opportune moment, improves the change a person will exhibit this behaviour. This reminder design is in line with earlier work regarding interruptibility.

H3: If COM-B reminders are given, opportunity mediates adherence. It was hypothesized that when people receive these automatic reminders, it was an opportune moment to perform the targeted behaviour. Therefore, it was expected that opportuneness was associated with people's adherence, thereby partly explaining an increase in adherence for COM-B reminders.

4.3. Method

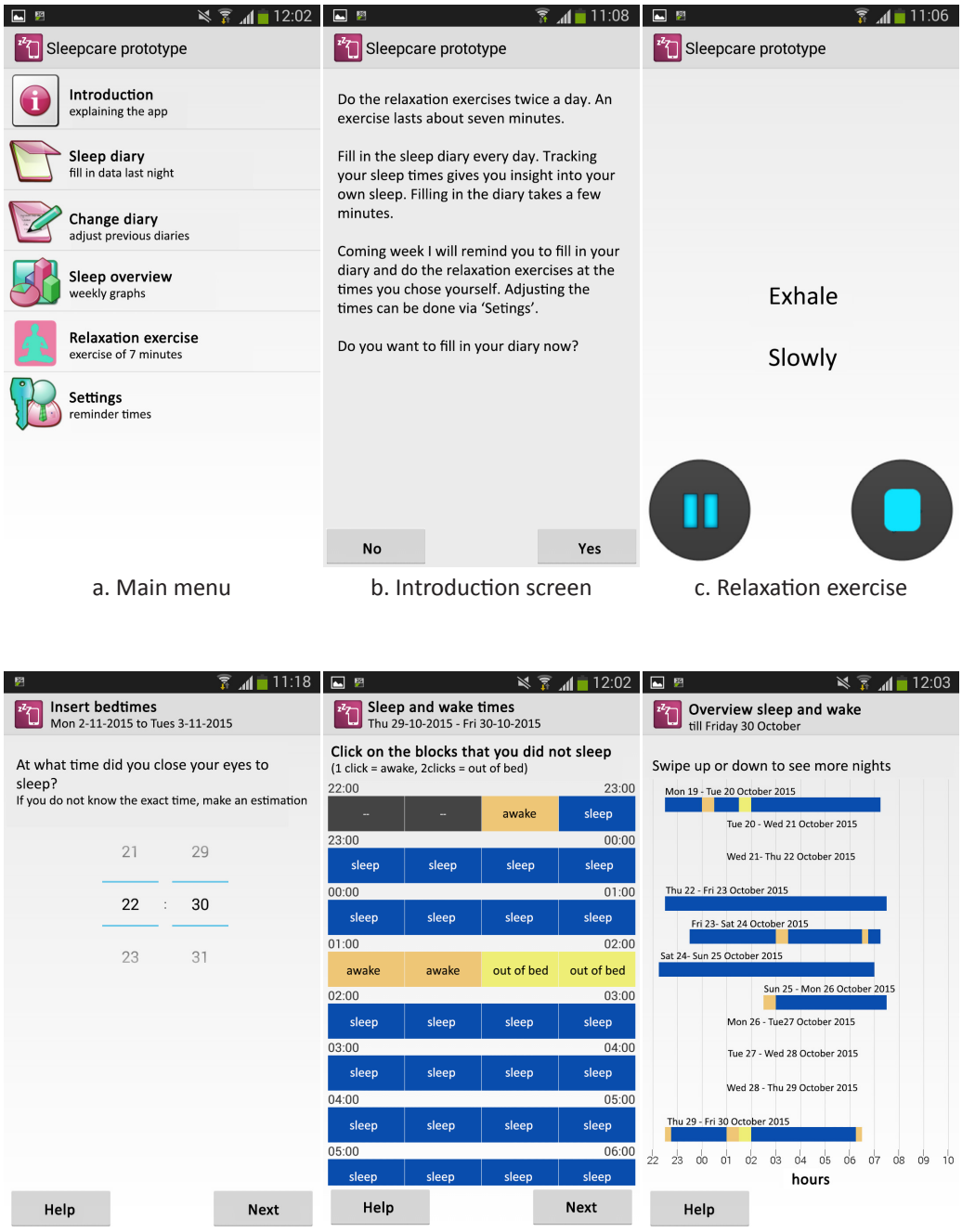
4.3.1. Experimental design

This field experiment had a within-subjects design with 45 participants who were exposed to three conditions during a total time of three weeks. In one condition participants received no reminders to perform targeted behaviour, in the other condition participants set the reminders themselves, and the last condition consisted of automatic COM-B reminders. The order of the three conditions was counter-balanced across the participants. The study was approved by Human Research Ethics Committee of Delft University of Technology.

4.3.2. The intervention system

An app for people suffering from insomnia was developed to test the two types of reminders. Since most people always carry their phones with them, smartphones were suitable for reminders. The app contained a sleep diary, a relaxation exercise, sleep overview graphs, and reminders. The two different activities were chosen to measure adherence because they have different properties. For example, it was probably easier for people to spend 1 minute, which is the approximated time for filling in the diary, than 7 minutes, which is the length of the relaxation exercise.

Navigation in the app was done using the main menu (Figure 2a), containing all the elements of the app. In the introduction screen (Figure 2b) a short explanation of the relaxation exercise and the sleep diary was given, as well as information about the reminders. Furthermore, the app contained a progressive muscle relaxation exercise (Figure 2c). The instructions were both visual on the screen in text and simultaneously audible. Moreover, the app contained a sleep diary consisting of four screens each with one question on it, respectively: (1) what time did you go to sleep?, (2) what time



a. Main menu b. Introduction screen c. Relaxation exercise

d. Sleep diary question e. Diary graphical question f. Sleep diary overview

Figure 2. Screenshots of the Sleepcare app translated from Dutch.

did you get up?, (3) indicate when you were awake, (4) which score would you give your sleep? (Figure 2d and 2e). Via a different menu-item users could correct mistakes in their sleep diaries. An overview of users' tracked sleep was shown in a graphical overview (Figure 2f). The last menu-item was settings. Here, people could change the times of the reminders in the self-set condition. Participants chose the time for three daily reminders: one reminder for filling out their sleep diary, and two others for doing the relaxation exercise.

For the automated reminders, we assumed that if people download an app, they were able and motivated to use the app. Leaving one requirement to exhibit a behaviour to fulfil, namely opportunity. The opportunities to exhibit the targeted behaviour were automatically detected by the smartphone based on the smartphone usage. The sleep diary needs to be filled out as closely as possible to waking up. So, users received a reminder for filling in the diary the first time they turned on their phone in the morning. Reminders for the relaxation exercises were generated when people: a) were at the same physical place for more than one hour, since it was regarded appropriate to take a break when sitting still for some time; b) ended a phone call, and thus were already distracted from another (important) task; or c) used another app1, and thus were already using their phone. Reminders were of course not generated when users already filled in their diary or did the relaxation exercise, or were occupied with that activity. Moreover, the time between two reminders was at least 30 minutes.

4.3.3. Procedure

Participants were recruited via social media, online advertisements, the website www.ikgalekkerslapen.nl, and in university lectures. After giving online consent and completing the first questionnaire, participants were enabled to download the app. The Sleepcare app ran on Android OS versions from 2.3 and higher. Participants used their own smartphone with Internet connection for this experiment. A short introduction text about the app was shown the first time participants opened the app (Figure 2b). The participants were instructed to fill out the sleep diary every day, and do the relaxation exercise twice a day. However, they were free to use the app in whatever way they wanted. After every week, participants got an e-mail with a link to that week's questionnaire and to inform them about the change of condition (within-subjects). Within two weeks after finishing the whole experiment participants were contacted for a semi-structured telephone interview.

People who gave informed consent, but did not download the app, received an e-mail with the question why they did not proceed with the experiment or app. Also participants from whom only a few days of data was received, got a similar e-mail to gain insight in reasons for non-adherence.

[1] Only the top-10 apps in the Netherlands triggered a reminder: Whatsapp, Facebook, Candy Crush, Nu.nl, Twitter, NOS, Wordfeud, Minecraft, Wheres My Water, Ruzzle, Browser.

4.3.4. Measurements

Primary outcome and mediation measures. During the experiment, adherence was measured by logging how often the diary was filled out (0 – 7 days), and how often the relaxation exercise was performed. These were the primary outcome measures. The mediation measures were the level of perceived self-empowerment and the suitability of the timing of the reminders (opportunity). Both mediation measures were measured using a questionnaire specially designed for this study consisting of 7-point Likert scale statements (Appendix J). Both concepts were measured in the two reminder conditions to examine if these concepts were able to explain adherence rates. Furthermore, motivation and ability were measured to check the assumptions of the COM-B model. The COM-B model states that the possibility that people exhibit a behaviour depends on the opportunity, the person's motivation to exhibit that behaviour, and the capability of a person to exhibit that behaviour (Michie et al., 2011). Opportunity was measured as one of the mediators and was expected to vary across the study. Motivation and capability, on the other hand, were expected to be constant during the experiment, so they would not influence adherence. In order to check this assumption motivation was measured using the Situational Motivation Scale (SIMS) (Guay, Vallerand, & Blanchard, 2000), consisting of Intrinsic Motivation (IM), Identified Regulation (IR), External Regulation (ER), Amotivation (AM). And ability was measured in three categories: ability to use a smartphone (AUS), ability to use the diary (AUD), and ability to use the relaxation exercise (AUR), using 7-point Likert scale statements (e.g., "*I know how I can respond to notifications on my smartphone*").

Exploratory measures. In order to perform more detailed, exploratory analyses that fall outside the main focus of the paper, the following measurements were taken. Firstly, an earlier study showed that objective (logged) and subjective (self-reported) adherence deviate from each other (Horsch et al., 2015). Therefore, participants' own estimation about their adherence was explored. This subjective adherence (SA) was measured by the questions "*How often did you fill in the sleep diary last week?*", and "*How often did you do the relaxation exercise last week?*". Because both objective and subjective adherence were measured in this study these two concepts could be compared to each other and the reliability of participants' own estimation about their behaviour and adherence could be derived.

Behavioural intention (BI), locus of control (LoC), irritation (Irr) and appreciation, and easiness to use the app in daily life (ETI) were measured to be able to examine possible associations between these variables and adherence rates. The theory of planned behaviour states that behavioural intention predicts behaviour (Ajzen & Madden, 1986). Therefore, behavioural intention (BI) was measured using six questions (e.g., "*I will follow the instructions/advice from the app*"). Locus of control (LoC) was measured via a 18-item Dutch questionnaire (Hertog, 1992). A higher internal locus of control has been found to influence diary adherence in an online lifestyle diary (Blanson Henkemans et al., 2009). Irritation and appreciation were respectively measured with

four 7-point Likert scale statements, and assigning a grade between 1 and 10 for the different reminders and app components. Reminders that irritate people because they are disruptive, or reminders that are not appreciated most likely decrease adhere (Bickmore, Mauer, Crespo, & Brown, 2007; Fischer et al., 2011; Smith et al., 2014).

Similarly, if an activity is hard to integrate in people's daily life, the probability that people will perform the activity decreases, since people's behaviour are affected by the principle of least effort (Zipf, 1949). Therefore, easiness to use was measured with six 7-point Likert scale statements. Furthermore, the Unified Theory of Acceptance and Use of Technology (UTAUT) measures technology acceptance and relates acceptance to usage (Venkatesh, Morris, Davis, & Davis, 2003), thereby possibly explaining adherence. Moreover, to investigate the possible effect of this minimal prototype on sleep the Insomnia Severity Index (ISI) was used (Bastien, Vallières, & Morin, 2001). See Appendices G - L for more details about these measures.

At the end of the experiment participants were contacted for a semi-structured telephone interview to explore their reasons for (none-)adherence. The subjects of the questions were: why people used the app, what their opinion was about the app and the separate parts of the app, how people used the app, if they noticed any effect (on sleep or in other ways), if there were any irritations, and if people had ideas for improvements or additions (See Appendix L for the used interview guide).

4.3.5. Participants

In total there were 45 participants who used the app for three weeks (Figure 3), 30 females and 15 males. The average age was 35 years (SD = 14). Their average ISI score was 13.5 (SD = 6.6), which is above the score of 10 (t = 2.60, p < .05) that is used as a cut-off for clinical levels of insomnia (Morin, 1993). The self-reported average of the ability to use a smartphone was 5.7 (SD = 1.3) on a scale from 1-7, which is an average rating on the positive side of the scale (t = 6.26, p < .01).

Although no strong conclusions can be drawn from a relatively arbitrary comparison with of the middle of the scale, participants seem positive about their smartphone abilities. Intention to use the app was 6.5 (SD = 0.7) on a score from 1-7, which also is an average rating on the positive side of the scale (t = 17.25 , p <.01). These scores suggested an adequate level of ability and motivation. As shown in Figure 3, 143 participants filled out the online informed consent form and the pre-measurements and 87 participants downloaded the app. From all those 87 participants automatic log files were received which describe their app usage (behaviour data). However, not everyone filled in the weekly online questionnaires. Sixteen participants filled in all three questionnaires, whereas 8 participants only filled in two questionnaires. The data of these 24 participants was used for the analyses regarding hypotheses two and three. Twenty-one participants filled in less than two questionnaires. Their logged behavioural

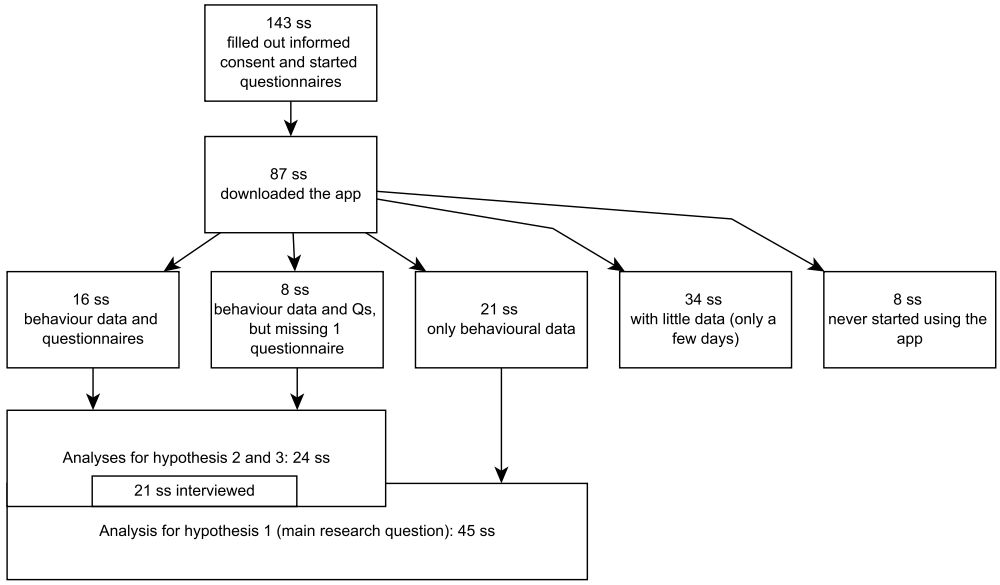


Figure 3. Flow diagram of participants in the experiment.
ss = participants, Qs = Questionnaires.

data, together with the data from the 24 previously mentioned participants, was used for analysis regarding hypothesis one. Thirty-four participants only used the app for a few days, and 8 participants did not use the app at all. As a consequence their data was not used for the analyses.

4.3.6. Statistical Analyses and Data Preparation

For data management and analyses, SPSS version 22 was used. To test hypothesis 1 Friendman's ANOVA tests were performed. Furthermore, posthoc Wilcoxon's test were done to investigate the differences between the separate conditions. Hypotheses 2 and 3 were tested by repeated measures mediation analyses. The analyses were done in line with the method described by Judd et al. (2001). The first step in this method is to test the overall treatment effects for the dependent variable (adherence), as well as for the mediators (self-empowerment and opportunity). This means testing for a difference in adherence between the three conditions using Friedman's ANOVAs (the same as used to test hypothesis 1), and testing for a difference in self-empowerment between the two reminder conditions using Wilcoxon's test, and testing for a difference in opportunity between the two reminder conditions using Wilcoxon's test. Although the original procedure suggests not to proceed with mediation analysis if no differences are found in this first step, Zhao et al. (2010) claim this is too strict. Zhao explains that mediation can occur even when there is no total effect of independent variable on the dependent variable, so the guidelines of Zhao et al. (2010) were followed. The second step was estimating the regression equation given by Judd et al. (2001) that indicates

$$Y_{Di} = (\delta_2 - \delta_1) + 0.5\,(\delta_{22} - \delta_{11})\,X_{Si} + 0.5\,(\delta_{22} + \delta_{11})\,X_{Di} \qquad (1)$$

Y_{Di}: Difference in diary/relaxation adherence between COM-B and Self-set condition
X_{Si}: Sum of self-empowerment/opportunity score in COM-B and Self-set condition
X_{Di}: Difference between self-empowerment/opportunity score in the COM-B and Self-set condition

Mediation is suggested if the 3rd coefficient ($0.5\,(\delta_{22}+\delta_{11})$) is significant.
Moderation is suggested if the 2nd coefficient ($0.5\,(\delta_{22}-\delta_{11})$) is significant.

Box 1. Regression equation belonging to repeated measure mediation (Judd et al., 2001)

mediator and moderator effects (Box 1). Four different regression equations were used with a bootstrapping samples of 1000. Each analysis took as dependent measure either adherence to the diary or relaxation exercises. Self-empowerment (SE) score or Opportunity (Opp) score were included as mediator in these analyses.

For testing hypothesis 1 the behavioural data collected via the app of 45 participants was used (Figure 3). For testing hypotheses 2 and 3 the questionnaire data was needed. We hypothesised that participants just forgot to fill in a weekly questionnaire unrelated to the week and condition, so that the data was missing completely at random. Little's MCAR test confirmed this assumption (Chi-square < 0.001, $df = 1692$, $p > 0.99$). Therefore, hypotheses 2 and 3 were tested using the data of the 16 complete datasets plus 8 datasets which missed one weekly questionnaire (amount of missing values: 388/5064 = 8% data points). This missing data was filled in using the expectation maximization method (Pigott, 2001) using all variables except the demographics.

4.3.6.1. Exploratory analyses

To explore a possible effect of the app on insomnia a t-test with the ISI scores measured before and after the experiment was performed. Furthermore, Cronbach's alphas were calculated for the concepts that were measured with multiple questions, such as opportunity (Opp), self-empowerment, and irritation (Irr). Items that affected Cronbach's alphas negatively were deleted resulting in Cronbach's alphas ranging from .43 to .99 (see Appendices G - K). The average scores for each concept were calculated with the remaining items. Next, Friedman's and Wilcoxon's tests were performed to test differences between conditions for repeated measures. Besides, correlations between adherence and repeatedly measured variables were calculated using the procedure of Bland and Altman (1995).

The interviews were analysed in line with the method of thematic analysis (Braun & Clarke, 2006). The first author (CH) familiarized herself with the data by conducting the interviews and transcribing the audio files. The data of all the participants was then organized per question, and codes were added to the answers. Lastly, the codes were grouped together in themes and a brief summary of the general gist was written.

Table 1. Wilcoxon tests with Bonferroni correction showing differences between no reminder and a reminder, but not between the two types of reminders.

Number of filled in diaries:	
No Reminder (*Mdn* = 4) vs. Self-Set (*Mdn* = 6)	*T* = 95.5, *p* = .002, *r* = -.27
No Reminder (*Mdn* = 4) vs. COM-B (*Mdn* = 7)	*T* = 67.0, *p* < .001, *r* = -.31
Self-Set (*Mdn* = 6) vs. COM-B (*Mdn* = 7)	*T* = 152.0, *p* =.78, *r* = -.02
Number relaxation exercises done:	
No Reminder (*Mdn* = 0) vs. Self-Set (*Mdn* = 1)	*T* = 42.5, *p* = .001, *r* = -.28
No Reminder (*Mdn* = 0) vs. COM-B (*Mdn* = 1)	*T* = 84.0, *p* = .011, *r* = -.22
Self-Set (*Mdn* = 1) vs. COM-B (*Mdn* = 1)	*T* = 192.5, *p* = .81, *r* = -.02

4.4. Results

4.4.1. Hypothesis 1 – the effect of reminders

The results confirmed hypothesis 1. In the no reminder condition a median of 4 (IQR = 6) filled in diaries per week was found, and a median of 0 (IQR = 6) performed relaxation exercises were done. In the self-set reminder condition in median of 6 (IQR = 2) was found for the diaries, and a median of 1 (IQR = 3) for the relaxation exercise. In the COM-B condition a median of 7 (IQR = 3) filled in diaries was found, and a median of 1 (IQR = 5) for the relaxation exercises (Table 1). Friedman's ANOVA's showed differences between the conditions for the number of diaries filled in (χ^2 (2) = 14.63, p = .001), and for the number of relaxation exercises done (χ^2 (2) = 9.04, p = .011). To further investigate the differences, Wilcoxon tests were performed in which the p-values were tested against Bonferroni corrected α-level of .0167. These analyses showed a difference between the condition without reminders and the conditions with reminders, but no differences were found between the two reminder conditions (Table 1).

4.4.2. Hypothesis 2 and hypothesis 3 – mediation

Hypotheses two and three were confirmed by repeated measures mediation analyses. The first step of Judd's (2001) procedure contains difference tests for the dependent variable and the mediators. The analyses done for hypothesis 1 already showed a difference in adherence (dependent variable) between the conditions. The differences for the mediators are shown in Table 2. Self-empowerment scores were on average higher in the self-set condition than in the COM-B condition (Table 3). The opportunity scores were on average higher in the self-set condition than in the COM-B condition for the relaxation exercise, for the diary this difference was not significant (Table 2). The second step of the repeated measures mediation analyses contains the four regression

Table 2. Means and standard deviations of the repeated measures, the difference between conditions for these measures, and their correlations across conditions with behavioural adherence.

Diary	NR *M* (*SD*)	Self-Set *M* (*SD*)	COM-B[d] *M* (*SD*)	Pearson's correlation
Self-Empowerment		4.72 (0.86)	3.84 (0.78)**	.50*
Opportunity		5.46 (1.17)	5.33 (1.46)	*.45**
Subjective Adherence	6.21 (1.76)	6.56 (1.01)	6.26 (1.59)	.67**
Satisfaction with adherence	5.89 (1.62)	6.22 (1.32)	6.19 (1.43)	.57**
Easy to initiate	5.65 (1.24)	5.67 (1.10)	5.94 (1.08)	.33*
Irritation [a]		5.31 (1.01)	5.44 (1.31)	.62**
Intrinsic Motivation	3.88 (1.07)	4.39 (1.28)	4.52 (1.20)**	.30*
Identified Regulation	5.19 (1.17)	5.22 (1.06)	5.36 (0.99)	.25
External Regulation	3.19 (1.31)	3.34 (1.56)	3.53 (1.18)	.01
Amotivation	1.91 (1.32)	2.21 (1.08)	2.12 (1.39)	.27
Appreciation		5.78 (2.80)[b]	4.96 (2.29)[b]	
Relaxation				
Self-Empowerment		4.36 (1.06)	3.07 (0.89)**	-.18
Opportunity		3.70 (1.50)	2.98 (1.59)*	*.11*
Subjective Adherence	6.65 (5.31)	7.63 (5.09)	6.44 (4.14)	.82*
Satisfaction with adherence	3.84 (2.14)	4.18 (1.99)	3.48 (1.97)*	.18
Easy to initiate	3.79 (1.81)	3.96 (1.49)	3.45 (1.59)	.00
Irritation [a]		5.12 (1.36)	4.23 (1.79)*	-.13
Intrinsic Motivation	3.46 (1.01)	3.62 (1.12)	3.57 (1.20)	.28*
Identified Regulation	4.92 (1.37)	5.07 (1.27)	5.07 (1.32)	.19
External Regulation	3.23 (1.44)	3.19 (1.39)	3.35 (1.56)	.14
Amotivation	2.80 (1.55)	2.59 (1.26)	2.64 (1.33)	.09
Appreciation	4.74 (3.04)[c]	7.78 (2.04) [c]	4.22 (2.67)[c]**	

r = Pearson's correlation between measured variable and adherence
* p < .05, ** p< .01
[a] the lower the number, the higher the irritation
[b] appreciation of the reminder type measured from 1 to 10 after a condition,
[c] appreciation of the reminder type measured from 1 to 10 after the whole experiment
[d] significant difference between conditions are denoted by a * in this column

Table 3. Means and standard deviations of the measurements, and their correlations with behavioural adherence

Measures (n = 24)	Mean (SD)	Pearson's Correlations Diary	Pearson's Correlations Relax
Pre-measures			
AUS	5.71 (1.28)	-.20	-.43*
Behavioural intention	6.48 (0.70)	.62**	.30
Locus of Control[a]	7.08 (3.44)	.10	.05
Insomnia Severity	13.50 (6.60)	.18	-.15
Post-measures			
Appreciation diary[b]	8.09 (1.77)	.39	-
Appreciation relax[b]	5.74 (2.36)	-	.53**
Appreciation app[b]	7.31 (1.55)	.35	.32
AUD	6.63 (0.76)	.59**	-
AUR	5.46 (1.64)	-	-.04
Insomnia Severity	11.28 (6.03)	-.01	-.21
UTAUT-concepts (post)			
Utility	5.03 (1.26)	.37	.20
Effort	6.41 (0.58)	.39	.33
Social influence	2.30 (1.45)	-.15	-.13
Facilitating conditions	4.93 (1.48)	.21	.16
Attitude	5.91 (0.81)	.09	.31
Self-efficacy	6.73 (0.53)	.25	-.21
Anxiety	1.95 (1.12)	.27	.19
Trust	5.43 (1.19)	.36	.16
Behavioural intention	5.87 (0.81)	.27	.10

AUS = Ability to Use a Smartphone,
AUD = Ability to Use the Diary,
AUR = Ability to Use the Relaxation exercise
* $p < .05$, ** $p < .01$
[a] scale ranges from 0 to 18, higher scores mean higher external locus of control
[b] grade given by the participant for the indicated component measured on a scale from 1-10

Regression functions

Diary adherence (diary self-empowerment): $Y_{Di} = -2.39 + 0.14\ X_{Si} + 0.39^{*}\ X_{Di}$
Diary adherence (diary opportunity): $Y_{Di} = -3.17 + 0.32\ X_{Si} + 0.48^{**}\ X_{Di}$
Relaxation adherence (relaxation self-empowerment): $Y_{Di} = 0.31 - 0.05\ X_{Si} - 0.24\ X_{Di}$
Relaxation adherence (relaxation opportunity): $Y_{Di} = 2.22 - 0.35\ X_{Si} - 0.25\ X_{Di}$

* $p < .001$, 95% *CI*: 0.14 – 0.59, mediation
** $p = .045$, 95% *CI*: 0.03 – 0.95, mediation

Box 2. Regression equations belonging to the repeated mediation analyses

equations (Box 2). In the mediation analysis on the number of diary entries for self-empowerment (SE) we found a significant positive mediation effect, whereby increase in self-empowerment was associated with increase in adherence. The second mediation analysis on diary adherence included opportunity (Opp) score as mediator. The analysis found also a significant mediation effect. Here an increase in opportunity score coincides with increase in number of diary entries filled out. Two similar mediation analyses were conducted on the adherence of the relaxation exercise. No significant mediation effects were found.

The COM-B model assumes that participants are able and motivated to perform the targeted behaviour. To test these assumptions the scores for ability and motivation are reported. On a seven-point scale (values 1-7) general ability (AUS) to use a smartphone was 5.7, ability to fill in the diary was 6.6, and ability to do the relaxation exercise was 5.5. These high values seem to confirm our assumption about participants' capability. The identified regulation (IR) scores were the highest among the motivation scores, which suggests that participants were mostly using the app, because they wanted to use the app. Furthermore, the average amotivation (AM) scores were low. These values again seem to confirm our assumption about participants' motivation.

4.4.3. Exploratory analyses

Table 3 shows that four variables were associated with adherence. Diary adherence was correlated to behavioural intention and ability to use the diary (AUD). Relaxation adherence was correlated to the appreciation for the relaxation exercise and the general ability to use a smartphone (AUS). As can be seen the UTAUT concepts were not found to be correlated neither to diary nor relaxation adherence. There was a significant difference in sleep severity (ISI) before (M = 13.50, SD = 6.60) and after intervention (M = 11.28, SD = 6.03) $t(23) = 2.74$, $p = .012$, $r = .50$. Although people suffer less from insomnia after using the app, it was not a clinically meaningful difference, which is set on a 6-point reduction (Yang, Morin, Schaefer, & Wallenstein, 2009).

Table 2 shows the variables that were measured repeatedly. Self-empowerment (SE) differed over the conditions for both the diary and the relaxation exercise. The opportunity score (Opp) only differed for the relaxation exercise, not for the diary. In addition, there was a correlation between self-empowerment and diary adherence, and between opportunity and diary adherence. These correlations were not found for the relaxation exercise.

4.4.4. Drop-outs

Thirty people responded to the question why they did not download the app or used it very little. The main reasons were a) unsuccessful in downloading the app, b) problems with the technology, c) inappropriate timing, and d) other reasons. The most prevalent problem was downloading the app. The app was provided via the Google Playstore as a test version, meaning participants had to become part of a Google group, as a result, people had to perform extra steps, which caused problems for people. Furthermore people experienced problems with the technology, e.g., their Android version was too old, or their smartphone broke. Besides technology-related problems, people mentioned that the timing for using the app was not convenient because they were, for example, rehousing or on holiday. Other reasons for dropping out were that people found another solution for their sleeping problem, they did not notice an effect, or they simply forgot to use it.

4.4.5. Interviews

The interviews indicated that most people were positive about the sleep diary *"[about the diary] It just worked well, it was crisp and clear, I did not have any problems.*" (female, 35 yr). In contrast to the relaxation exercise, which induced more diverse opinions. Some people had a positive attitude towards the relaxation exercise *" I was surprised that such an easy relaxation exercise helped me that much. I just had to do it every day."* (female, 39 yr), others thought the exercise was boring *"The relaxation exercise was so-so, especially because every time it was the exact same exercise, so after three days I was bored with it."* (female, 34 yr), and others prefered to do their own relaxation exercises with which they were already familiar *"I only did the relaxation exercise once or twice, because I already do breathing and meditation exercises. So, the relaxation exercise in the app didn't have any added value."* (male, 27 yr). About the reminders participants said that the COM-B reminders were annoying and that the timing was bad *"Well, the reminders came randomly, and then I experienced them as bothersome."* (female, 54 yr). In general, the self-set reminders were perceived as timed better *"I have the impression that the self-set reminders worked best for me. Those reminders came at the right moments."* (female, 20 yr), although a few people thought differently *"The self-set reminders were actually not much better than the automatic ones. Both came often at inconvenient times."*(female, 56 yr). In case people do not get any reminders, they just forget to do an activity *"It was inconvenient when I did not get*

a reminder, because then I forget to do the activities." (female, 34 yr). Interesting was that some people were waiting for the reminder to arrive and perform the activity, even in the No Reminder condition *"When I did not get any reminders, I was kind of waiting for them"* (female, 56 yr).

4.5. Discussion and conclusion

In this paper, we tested reminders in a mobile sleep intervention. On average, participants filled out the sleep diary more often with reminders than without reminders. Also, the relaxation exercise was performed more frequently with reminders compared to the no reminder condition. Both reminders increased adherence thereby supporting the first hypothesis. The results showed that it did not matter which kind of reminder participants received. Support for hypotheses two and three was also obtained, as we found significant mediating effects of self-empowerment and opportunity on adherence for the sleep diary in the regression analysis. However, no support for hypothesis two and three were found regarding mediation effects of self-empowerment and opportunity on adherence to the relaxation exercise. The results of the regression analyses showed a partly mediation, this means that for the self-set reminders, the associated feeling of self-empowerment can explain part of the diary adherence. For the COM-B reminder, one explaining factor is people's perception that the reminders were given at opportune moments. Besides self-empowerment and opportuneness, different mechanisms are likely at play to why people adhere to the reminders.

The findings show that perceived self-empowerment was higher in the self-set reminder condition than in the COM-B reminder condition, as expected. Opportuneness of the reminders between the two conditions only differed for the relaxation exercise (in opposite direction to expectation), but was not found for the diary entries. The lack in difference in opportuneness has probably been caused by the actual timing of the reminders for the diary, which did not differ that much between the two conditions. In the COM-B condition a diary reminder was sent the first time someone turned on their phone which is most likely shortly after they wake-up, in the self-set condition people probably set the diary reminder a short time after they wake-up as well. So, timing for the diary reminder in the two conditions were most likely very similar.

Exploratory analyses provided more insights in which cases underlying principles, such as self-empowerment and opportuneness, play a role. From the interviews we learned that participants had a negative attitude towards the relaxation exercise. This observation was supported by the relative low appreciation scores given to the relaxation exercise (5.7 on a scale from 1 to 10). This suggests that a positive attitude towards the activity might be a precondition for factors such as self-empowerment and opportuneness to come into play. In case of a negative attitude, which is the case for the relaxation

exercise, self-empowerment and opportunity did not explain adherence. A negative attitude probably deters people from exhibiting the targeted behaviour, irrespective of the level of perceived self-empowerment or opportuneness of the moment. Therefore, another sort of trigger might be more suitable for the relaxation exercise.

Several mechanisms have been suggested for why people adhere to reminders. Fogg describes three types of triggers in his behavioural model (Fogg, 2009). If we apply Fogg's categorization of triggers, the reminders in this experiment mostly resemble signals. We speculate however that the relaxation exercise would benefit more from sparks than from signals, since the appreciation for the relaxation exercise was low. According to Fogg (2009) there are three core motivators that sparks can use: pleasure-pain, hope-fear, and social acceptance-rejection. For the relaxation exercise the reminder could for example emphasize the relaxed state people experience (pleasure) while doing the relaxation exercise. Future research could explore the effect of these different types of reminders.

When examining adherence, it is important to study the participants who dropped-out. By studying the drop-outs insight can be gained about the underlying reasons for not doing something. Approximately half of the participants who downloaded the app only used it for a few days or even did not use it at all. We did our best to contact those people and discover their reasons, which were mainly technical problems, and inappropriate timing to participate in the study. The possible implications of these drop-outs for our results are unknown. It might be the case that more persevere people, or people that already heavily use their phone participated longer in the experiment. Apart from drop-outs, increasing experimental compliance (e.g., filling in weekly questionnaires) also requires attention to obtain the required data set, especially in experiments in the field. In this study approximately 50% of the participants who used the app filled in the questionnaires. Therefore not all participants could be included in the analysis, and some missing data was estimated. Nonetheless, field studies are necessary to ecologically validate mobile interventions, and irreplaceable when studying adherence.

To fully appreciate the findings, it is important to consider the study's limitations. The main limitation of this study is the implementation of the COM-B reminders. A relative simple algorithm was implemented to detect opportune moments to perform the target behaviour. However, as mentioned before, this might have resulted in diary reminders to occur at similar moments in the two reminder conditions. Furthermore, the algorithm did not anticipate on participants who use their phone minimally. For example, participants might not have received COM-B reminders, if they did not use WiFi. Future research might therefore explore ways to improve the algorithm. Another limitation is the extent of the intervention system. Applications that offer more support, such as cognitive therapy or sleep restriction, might elicit more positive usage attitude. Adherence to reminders might be higher in these applications. On the other hand, applications that offer little support to which people have negative attitudes

might also benefit from reminders. For example, adherence to mundane tasks such as hour registration, might improve due to reminders. Next to self-empowerment and opportuneness, other underlying principles, like obligation to employers, probably play a role in such processes.

The main contribution of this study can be summarized by two new insights. First, the study shows that reminders do improve adherence to target behaviours such as keeping a sleep diary and performing relaxation exercises. This is important as adherence has been associated with treatment effect (Horsch et al., 2015). Second, self-empowerment and opportunity can partly explain why people follow up on reminders and perform the desired activity.

4.6. References

Ajzen, I., & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. Journal of Experimental Social Psychology, 22(5), 453-474.

Andersson, G. (2010). The promise and pitfalls of the internet for cognitive behavioral therapy. BMC Medicine, 8. doi: 10.1186/1741-7015-8-82

Armstrong, A. W., Watson, A. J., Makredes, M., Frangos, J. E., Kimball, A. B., & Kvedar, J. C. (2009). Text-message reminders to improve sunscreen use: a randomized, controlled trial using electronic monitoring. Archives of dermatology, 145(11), 1230-1236.

Bastien, C. H., Vallières, A., & Morin, C. M. (2001). Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep medicine, 2(4), 297-307.

Bentley, F., & Tollmar, K. (2013). The Power of Mobile Notifications to Increase Wellbeing Logging Behavior. Paper presented at the ACM SIGCHI International Conference on Human Factors in Computing Systems.

Bickmore, T., Mauer, D., Crespo, F., & Brown, T. (2007). Persuasion, task interruption and health regimen adherence. Paper presented at the Persuasive Technology.

Bland, J. M., & Altman, D. G. (1995). Calculating correlation coefficients with repeated observations: Part 2—Correlation between subjects. Bmj, 310(6980), 633.

Blanson Henkemans, O. A., van der Boog, P. J. M., Lindenberg, J., van der Mast, C. A. P. G., Neerincx, M. A., & Zwetsloot-Schonk, B. J. H. M. (2009). An online lifestyle diary with a persuasive computer assistant providing feedback on self-management. Technology and Health Care, 17(3), 253-267.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative research in psychology, 3(2), 77-101. doi: 10.1191/1478088706qp063oa

Campbell, J. D., & Tesser, A. (1983). Motivational interpretations of hindsight bias: An individual difference analysis. Journal of personality, 51(4), 605-620.

Cheung, C., & Lee, M. K. (2000). Trust in Internet shopping: A proposed model and measurement instrument. AMCIS 2000 Proceedings, 406.

Deng, X., Wang, Y., Zhu, T., Zhang, W., Yin, Y., & Ye, L. (2015). Short Message Service (SMS) can Enhance Compliance and Reduce Cancellations in a Sedation Gastrointestinal Endoscopy Center: A Prospective Randomized Controlled Trial. Journal of medical systems, 39(1), 1-11.

Dimatteo, M. R., Giordani, P. J., Lepper, H. S., & Croghan, T. W. (2002). Patient adherence and medical treatment outcomes: a meta-analysis. Medical care, 40(9), 794-811.

Dismukes, R. K. (2012). Prospective memory in workplace and everyday situations. Current directions in psychological science, 21(4), 215-220.

Donkin, L., Christensen, H., Naismith, S. L., Neal, B., Hickie, I. B., & Glozier, N. (2011). A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res, 13(3).

Donkin, L., & Glozier, N. (2012). Motivators and Motivations to Persist With Online Psychological Interventions: A Qualitative Study of Treatment Completers. J Med Internet Res, 14(3).

Edinger, J. D., & Means, M. K. (2005). Cognitive-behavioral therapy for primary insomnia. [Review]. Clin Psychol Rev, 25(5), 539-558. doi: 10.1016/j.cpr.2005.04.003

Espie, C. A., MacMahon, K., Kelly, H., Broomfield, N. M., Douglas, N. J., Engleman, H. M., . . . Wilson, P. (2007). Randomized clinical effectiveness trial of nurse-administered small-group cognitive behavior therapy for persistent insomnia in general practice. SLEEP-NEW YORK THEN WESTCHESTER-, 30(5), 574.

Eysenbach, G. (2005). The law of attrition. J Med Internet Res, 7(1).

Fenerty, S. D., West, C., Davis, S. A., Kaplan, S. G., & Feldman, S. R. (2012). The effect of reminder systems on patients' adherence to treatment. Patient Prefer Adherence, 6, 127-135.

Festinger, L. (1957). A theory of cognitive dissonance. Evanston, IL: Row, Peterson.

Fischer, J. E., Greenhalgh, C., & Benford, S. (2011). Investigating episodes of mobile phone activity as indicators of opportune moments to deliver notifications. Paper presented at the Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services.

Fogg, B. J. (2009). A behavior model for persuasive design. Paper presented at the Persuasive 2009, Claremont, California, USA.

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: an integrated model. MIS quarterly, 27(1), 51-90.

Gould, R. A., & Clum, G. A. (1993). A meta-analysis of self-help treatment approaches. Clinical Psychology Review, 13(2), 169-186.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). Motivation and emotion, 24(3), 175-213.

Hertog, P. C. d. (1992). Instrumenteel onderzoek. De 'IE-18 locus of control'- vragenlijst: betrouwbaarheid en validiteit van een gewijzigde versie. Nederlands tijdschrift voor de psychologie, 47, 82-87.

Ho, J., & Intille, S. S. (2005). Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.

Horsch, C. H. G., Lancee, J., Beun, R. J., Neerincx, M. A., & Brinkman, W. P. (2015). Adherence to Technology-mediated Insomnia Treatment: A meta-analysis, interviews with users, and focus groups with users and experts. Journal of Medical Internet Research, 17(9), e214. doi: 10.2196/jmir.4115

Hudson, S., Fogarty, J., Atkeson, C., Avrahami, D., Forlizzi, J., Kiesler, S., . . . Yang, J. (2003). Predicting human interruptibility with sensors: a Wizard of Oz feasibility study. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.

Jones, R. K., Darroch, J. E., & Henshaw, S. K. (2002). Contraceptive use among US women having abortions in 2000-2001. Perspectives on sexual and reproductive health, 294-303.

Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. Psychological methods, 6(2), 115.

Kaltenthaler, E., & Cavanagh, K. (2010). Computerised cognitive behavioural therapy and its uses. Progress in Neurology and Psychiatry, 14(3), 22-29.

Kliegel, M., & Martin, M. (2003). Prospective memory research: Why is it relevant? International Journal of Psychology, 38(4), 193-194.

Krishna, S., Boren, S. A., & Balas, E. A. (2009). Healthcare via cell phones: a systematic review. Telemedicine and e-Health, 15(3), 231-240.

Krueger, K. P., Felkey, B. G., & Berger, B. A. (2003). Improving adherence and persistence: a review and assessment of interventions and description of steps toward a national adherence initiative. Journal of the American Pharmacists Association, 43(6), 668-679.

McDaniel, M. A., & Einstein, G. O. (2007). Prospective memory: An overview and synthesis of an emerging field: Sage Publications.

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implementation Science, 6(1), 42.

Morin, C. M. (1993). Insomnia: Psychological assessment and management: Guilford Press New York.

Ohayon, M. M. (2002). Epidemiology of insomnia: what we know and what we still need to learn. Sleep medicine reviews, 6(2), 97-111.

Pejovic, V., Musolesi, M., & Mehrotra, A. (2015). Investigating The Role of Task Engagement in Mobile Interruptibility. Paper presented at the Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct.

Pielot, M., De Oliveira, R., Kwak, H., & Oliver, N. (2014). Didn't you see my message?: predicting attentiveness to mobile instant messages. Paper presented at the Proceedings of the 32nd annual ACM conference on Human factors in computing systems.

Pigott, T. D. (2001). A review of methods for missing data. Educational research and evaluation, 7(4), 353-383.

Rosekind, M. R., & Gregory, K. B. (2010). Insomnia risks and costs: health, safety, and quality of life. Am J Manag Care, 16(8), 617-626.

Rothschild, J. M., Landrigan, C. P., Cronin, J. W., Kaushal, R., Lockley, S. W., Burdick, E., . . . Czeisler, C. A. (2005). The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care. Critical care medicine, 33(8), 1694-1700.*

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American psychologist, 55(1), 68.

Smith, J., Lavygina, A., Ma, J., Russo, A., & Dulay, N. (2014). Learning to recognise disruptive smartphone notifications. Paper presented at the Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services.

Smith, J. D., & Oakley, D. (2005). Why Do Women Miss Oral Contraceptive Pills? An Analysis of Women's Self-Described Reasons for Missed Pills. Journal of Midwifery & Women's Health, 50(5), 380-385.

Turner, L. D., Allen, S. M., & Whitaker, R. M. (2015). Interruptibility prediction for ubiquitous systems: conventions and new directions from a growing field. Paper presented at the Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

Unni, E. J., & Farris, K. B. (2011). Unintentional non-adherence and belief in medicines in older adults. Patient Educ Couns, 83(2), 265-268.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.

Vidal, C., Garcia, M., Benito, L., Milà, N., Binefa, G., & Moreno, V. (2014). Use of text-message reminders to improve participation in a population-based breast cancer screening program. Journal of medical systems, 38(9), 1-7.

Wantland, D. J., Portillo, C. J., Holzemer, W. L., Slaughter, R., & McGhee, E. M. (2004). The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res, 6(4).

World Health Organization, W. (2003). Adherence to long-term therapies: evidence for action. Geneva, Switzerland: World Health Organization.

Yang, M., Morin, C. M., Schaefer, K., & Wallenstein, G. V. (2009). Interpreting score differences in the Insomnia Severity Index: using health-related outcomes to define the minimally important difference. Current Medical Research & Opinion, 25(10), 2487-2494.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. Journal of consumer research, 37(2), 197-206.

Zipf, G. K. (1949). Human behavior and the principle of least effort. Cambridge, Massachusetts: Addison-Wesley Press.

5. Smartphone delivered cognitive behavioral treatment for insomnia: A randomized wait-list controlled trial

Abstract

Objective - To investigate the efficacy of Cognitive Behavioural Therapy for Insomnia (CBT-I) delivered via a smartphone app, compared to a waiting list control group.

Methods - In a randomised controlled trial a CBT-I app was compared to a waiting list condition. Participants with relatively mild insomnia disorder were recruited and randomly allocated to the app (n = 74) or the waiting list condition (n = 77). The app packaged a sleep diary, a relaxation exercise, sleep restriction, and sleep hygiene and education. The app was fully automated and adjusted itself to a participant's progress. The programme spanned six to seven weeks, after which participants received post-test measurements and a three-month follow-up. The participants in the waiting list condition received the app after they completed the post-test questionnaire. The measurements consisted of questionnaires and seven-day online diaries. The questionnaires measured insomnia severity, dysfunctional beliefs about sleep, and anxiety and depression symptoms. The diary measured sleep variables such as sleep efficiency. Multilevel analyses were performed to study the interaction effects between time and condition.

Results - The results show significant interaction effects (p < .01) favouring the app condition on the primary outcome measures of insomnia severity (d = -.66) and sleep efficiency (d = .71). Overall these improvements were also retained in a three-month follow-up.

Conclusion - This study demonstrated the efficacy of a fully automated smartphone app in the treatment of relatively mild insomnia. The effects were in the range of what is found for online treatment in general. This supports the applicability of such technical tools in the treatment of insomnia.

Chapter submitted as:

Horsch, C. H. G., Lancee, J., Griffioen-Both, F., Spruit, A. G. L., Fitrianie, S., Neerincx, M. A., Beun, R. J., Brinkman, W. P. (submitted). Smartphone delivered cognitive behavioral treatment for insomnia: A randomized wait-list controlled trial. *Journal of Medical Internet Research*.

5.1. Introduction

Approximately 10% of the general population suffers from chronic insomnia (Ohayon, 2002)mensen die zichzelf aanwijzen als poor sleepers, have insomnia problems. People with insomnia experience difficulties falling asleep and/or staying asleep, and as a consequence they suffer from sleep deprivation during the day (American Psychiatric Association, 2013). For example, insomnia is associated with low concentration, fatigue, and impaired cognitive functioning (Espie, Kyle, Hames, Cyhlarova, & Benzeval, 2012; Kyle, Morgan, & Espie, 2010; LeBlanc et al., 2007). Another consequence of insomnia is an increased risk of developing mental disorders like depression and anxiety (Baglioni et al., 2011; Taylor, Lichstein, Durrence, Reidel, & Bush, 2005), or physical complaints such as diabetes and high blood pressure (Mallon, Broman, & Hetta, 2005; Suka, Yoshida, & Sugimori, 2003). Insomnia also leads to societal costs such as reduced productivity, higher levels of sick-leave, and more accidents (Daley, Morin, LeBlanc, Grégoire, & Savard, 2009).

One of the most common nonpharmacological treatments used to treat insomnia is cognitive behavioural therapy (CBT-I). CBT-I is an effective treatment in either a face-to-face (Irwin, Cole, & Nicassio, 2006; Morin et al., 2006; Morin et al., 1999) or a self-help format (Ho et al., 2014; van Straten & Cuijpers, 2009). Recently it has become more common to offer these self-help formats via the internet. In a very recent meta-analysis (Zachariae, Lyby, Ritterband, & O'Toole, 2016) it was demonstrated that internet-delivered CBT-I showed large treatment effects (Cohen's d = 1.0) on the Insomnia Severity Index (ISI). In addition to the efficacy of computerised CBT-I (CCBT-I), it has multiple other advantages compared to face-to-face treatments. Potentially it can save costs, because less therapist time is needed and the treatment can be offered to a larger number of people who can go through the treatment in their own time.

Until now, studies on computerised treatments have been mostly limited to online treatments. A possible next step is delivering CCBT-I via a smartphone app. CCBT-I delivered via a mobile phone has similar advantages to existing CCBT-I, such as wide and easy accessibility, reducing stigma, and cost-efficiency (Voinescu, Szentagotai, & David, 2013), but it could potentially exceed those advantages since smartphones are portable. People carry their phones with them all the time and they are ubiquitous, unobtrusive, and intimate. Therefore, an effective app-based treatment for insomnia would increase the possible coverage for CBT-I. Furthermore, smartphones are rich in sensors, computationally powerful, and remotely accessible, which provides opportunities for personalisation, ecological momentary access, and real-time tracking (Kaltenthaler & Cavanagh, 2010; Konrath & Yan, 2015).

In the domain of sleep, several kinds of sleep apps have been studied. For example, there is an app that unobtrusively increases awareness of sleep hygiene recommendations (Bauer et al., 2012), an app that applies active sleep sampling for measuring sleep (Lawson et al., 2013), and a social app that shares time in bed based on alarm usage (Shirazi et al., 2013). To our best knowledge, however, there are no studies that evaluate

the efficacy of standalone CBT-I apps. To bridge this gap, a randomised controlled trial (RCT) was conducted to compare a CBT-I based app with a waiting list control group. It was expected that this app would have an ameliorating effect on insomnia severity and sleep impairment compared with the waiting list control group, assessed by a sleep diary.

The app offered a sleep diary, a relaxation exercise, sleep restriction, and sleep hygiene and education. Since sleep restriction is seen as the most effective exercise (Harvey, Inglis, & Espie, 2002; Miller et al., 2014), it was the main focus of the app. The goal was to demonstrate the app's efficacy in a sample of patients with relatively mild insomnia in order to test the proof of principle before investigating it in more a severely affected population.

5.2. Method

This study had a between-subjects design with two arms: a waiting list condition and an intervention condition with pre-intervention, post-intervention, and three-month follow-up measures.

5.2.1. Participants

Participants were recruited from 15th of August to 21st of October 2015 via websites, social media, online advertisements, flyers, and a press release. An initial group of 640 interested individuals completed an informed consent form and started the online questionnaire. Of this group, 269 people were excluded based on the inclusion/ exclusion criteria (see Figure 1). Inclusion criteria were: a) difficulty with initiating or maintaining sleep for at least thirty minutes a night, for at least three nights a week, for at least three months, causing clinically significant distress and/or impairment in daily functioning, i.e. meeting the criteria for a DSM-5 diagnosis of insomnia (DSM-5); b) stable medication; c) aged eighteen years or older; and d) a valid email address, connected to the internet, and in possession of an Android smartphone (operating system version 4.1 or higher). Exclusion criteria were: a) total sleep time ≤ five hours on average as reported in a consecutive seven-day sleep diary prior to the experiment; b) Insomnia Severity Index score lower than 7; c) earlier treatment with CBT-I; d) started other psychotherapy in the last six months; e) self-reported diagnosis of schizophrenia or psychosis; f) alcohol or marijuana abuse (more than three glasses of alcohol a day for at least twenty-one days a month, or use marijuana more than once a week); g) possible sleep apnoea (determined with a subscale of the SLEEP-50, cut off ≥ 15 (Spoormaker, Verbeek, van den Bout, & Klip, 2005)) ; h) shift-work; i) pregnant or breastfeeding; j) current suicidal plans; or k) symptoms of depression (determined with a subscale of

the CES-D, cut off ≥ 27). The final 151 participants were randomly assigned to the app (*n* = 74) or a waiting list condition (*n* = 79). Participants had a mean age of 39.66 years (*SD* = 13.44; range 18–80 years). Of the total 151 participants, 94 were female (62.3%). Demographic information about the participants can be found in Table 1.

5.2.2. Intervention

The Sleepcare app (Beun et al., Accepted 2016; Beun, Griffioen-Both, Ahn, Fitrianie, & Lancee, 2014) was based on previously published protocols (e.g. (Morin & Espie, 2003; Verbeek & Klip, 2005)). The app packaged a sleep diary, relaxation exercises, sleep restriction, and sleep hygiene and education (Figure 2). The app offered these exercises in Dutch, adjusted them to the participant, and reminded participants to perform the exercises. The basic programme spanned between six and seven weeks, depending on a participant's adherence. For example, if a participant had filled out fewer than six sleep diaries since using the app, the app explained to the participant that the sleep restriction exercise could only start after they had completed six diaries. If fewer than six days were completed, the introduction of the sleep restriction exercise was postponed until participants met this prerequisite. The app was fully automated and did not require any input from therapists or a human administrator. Automatic warnings were built in when participants slept for less than five hours on average. The first warning appeared after five days and warned against activities such as driving a car while feeling sleepy. Follow-up warnings also included a referral to the general practitioner, and the sleep restriction exercises were automatically stopped by the app.

The app consisted of a home screen on which the scheduled exercises for that current day were displayed. Furthermore, there was a menu, a calendar, and a conversation screen (Figure 3). The menu provided access to all components of the app and the CBT-I exercises. The calendar displayed all the scheduled activities for the whole seven weeks, which the participants could browse through at any time. The app interacted with the participants via dialogues on the conversation screen.

5.2.2.1. Conversations

The conversation screen displayed the dialogues between the app and the participants (Figure 3). Typically the app gave information and asked multiple-choice questions. In the conversations new exercises were introduced, and the progress of the participants was evaluated. Participants could only open new conversations after the previous conversation was finished. Conversations were initiated by the app based on a participant's adherence and progress. For example, if a relaxation exercise was done fewer than three times within four days of its introduction, a conversation would start to address the participant's adherence. Additionally, the app also started conversations based on a participant's progress – e.g. after filling in the sleep diary for a week the app calculated and showed weekly sleep efficiency averages in an evaluation conversation.

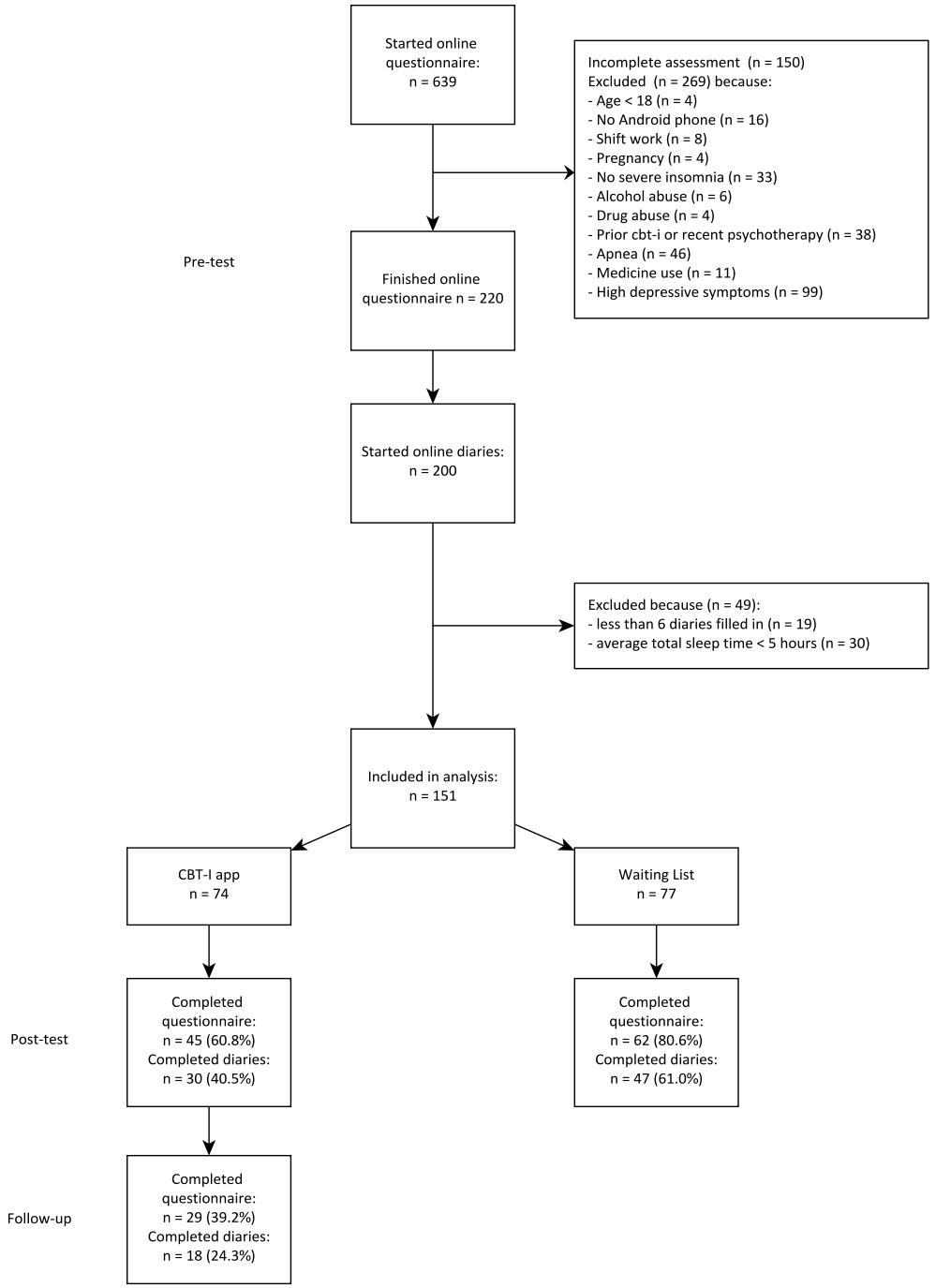


Figure 1. Flow diagram of the recruitment, reasons for exclusion, and experimental compliance

Table 1. Demographic per condition

		WL	App	
		M (*SD*)	*M* (*SD*)	
Age		41 (13.9)	39 (13.0)	$t(149) = 1.02$, $p = .31$
		n	*n*	
Gender	Female	49 (64%)	45 (61%)	χ^2 (1) = .13, p = .72
	Male	28 (36%)	29 (39%)	
Living together	Yes	49 (64%)	50 (68%)	χ^2 (1) = .26, p = .61
	No	28 (36%)	24 (32%)	
Currently employed	Yes	56 (73%)	58 (78%)	χ^2 (1) = .65, p = .42
	No	21 (27%)	16 (22%)	
Education	Lower general secondary education	7 (9%)	4 (5%)	χ^2 (3) = 1.10, p = .78
	Higher general secondary education	10 (13%)	9 (12%)	
	Community college	11 (14%)	9 (12%)	
	University	49 (64%)	52 (70%)	
Years insomnia	<1 year	8 (10%)	9 (12%)	χ^2 (4) = 5.40, p = .25
	1-5 years	38 (549%)	27 (36%)	
	5-10 years	10 (13%)	13 (18%)	
	>10 years	12 (16%)	20 (27%)	
	Unclear answer	9 (12%)	5 (7%)	
Insomnia due to a physical condition	Yes	9 (12%)	7 (9%)	χ^2 (1) = 0.20, p = .66
	No	68 (88%)	67 (91%)	
Prescribed Sleep medication	Yes	3 (100%)	6 (75%)	χ^2 (1) = 0.92, p = .35
	No	0 (0%)	2 (25%)	

5.2.2.2. CBT-I exercises

Sleep diary

The sleep diary was a visual translation of the core consensus sleep diary (CSD) (Carney et al., 2012) consisting of four screens asking participants to fill in their bed- and wake times, and their subjective sleep quality. The sleep diary could only be filled out for the previous night. Connected to the sleep diary was the sleep overview, which gave a visual summary of the participant's sleep.

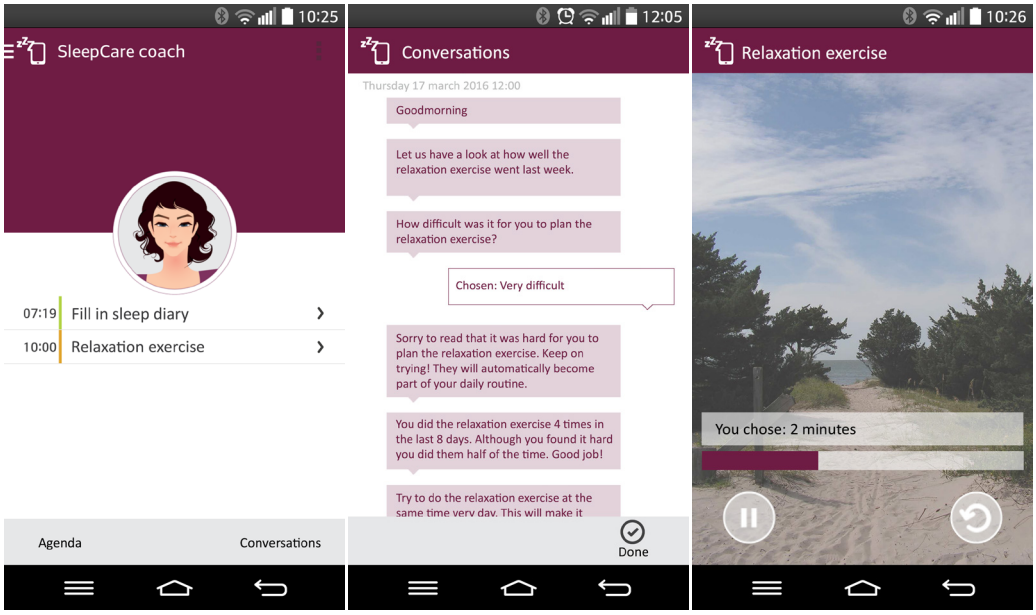
Figure 2. Treatment planning

Relaxation exercise

The relaxation exercise was a progressive muscle relaxation exercise of varying durations, ranging from one to sixteen minutes. The participants could choose the length of the exercise themselves before starting the exercise. By offering short exercises participants were able to gradually develop the habit of relaxing. The participants were guided by a voice track which told them which muscles to contract and when to relax. The app advised participants to do the relaxation exercise once a day, although participants could do the exercises as often as they wanted (Eijk, 2013).

Modified sleep restriction

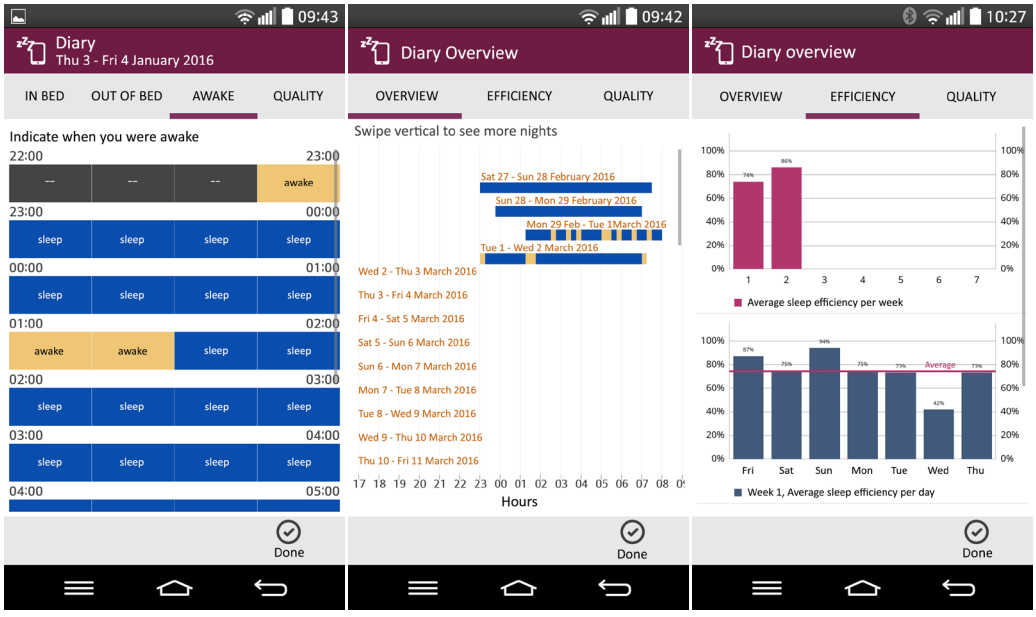
After a week the app introduced the modified sleep restriction exercise, on condition that participants filled out at least six sleep diaries and had an average sleep efficiency of less than 85%. An algorithm was developed to calculate the ideal and maximum time in bed for that specific participant based on the average sleep times of the previous week (Griffioen-Both et al., (in preparation)). The algorithm used rules such as: a) the ideal time in bed is equal to the average Time In Bed (TIB); b) the maximum time in bed is always at least one hour less than the average TIB; c) the advised time in bed is never less than five hours (Kyle et al., 2015); and d) the advised time in bed lies between average Total Sleep Time (TST) and average TIB. Participants had the opportunity to negotiate over their sleep time. The app first suggested that the participant abides by the ideal time in bed. The participant was then given the option to accept that time in bed, to negotiate longer time in bed up to the calculated maximum time in bed, or to refrain from sleep restriction at all. Participants were allowed to negotiate about and refrain from the sleep restriction exercise in order to enhance self-empowerment, set realistic goals, and thereby increase adherence. Every week the app evaluated the adherence and effect of the sleep restriction exercise. When participants' sleep efficiency exceeded 85%, they were allowed fifteen minutes extra in bed. When a participant's average sleep efficiency was lower than 85%, the app suggested the same or a further



Home screen with the scheduled exercises for this day

Conversation between app and participant

Relaxation exercise



Filling in the sleep diary (visually)

Overview of the sleep diary

Overview of the sleep efficiency per week and per day

Figure 3. Screenshots of the Sleepcare app translated from Dutch

restriction of fifteen minutes, depending on a participant's adherence. A modified sleep restriction protocol was used because we assumed this would increase the possibility of completing the restriction exercises (on the basis that more lenient sleep restriction is better than no sleep restriction at all).

Sleep hygiene and education

Sleep hygiene and education were presented on different screens as tips and facts in text format. The tips were divided over three categories: a) Food and Drink; b) Bedroom; and c) Behaviour. 'Use your bedroom only for sleeping, not for working' was an example of a Bedroom tip. The sleep facts were categorised into eight groups: a) Sleep cycles; b) Amount of sleep; c) Age; d) Animals; e) Disorders; f) Causes; g) Sleep medicine; and h) Fun facts. An example of a fact about Age was 'The amount of sleep a person needs is age dependent'.

5.2.2.3. Persuasive strategies

Different kinds of persuasive strategies were implemented to support the participants' adherence. First, the app sent notifications for both the exercises and the conversations. So, for a scheduled exercise such as filling in the sleep diary, the participants received a notification from the app. For unfinished conversations, additional reminders were sent every day at noon. Second, the app provided room for negotiation about the sleep restriction assignment as described earlier. Furthermore, the app was designed to be easy to use and attractive, to improve adherence (Beun et al., Accepted 2016; Consolvo, McDonald, & Landay, 2009).

5.2.3. Measurements

5.2.3.1. Questionnaire measures

Primary measure

Insomnia severity

The severity of insomnia was measured with a Dutch translation of the widely used Insomnia Severity Index (ISI). This is a seven-item questionnaire with scores ranging from 0 (no insomnia) to 28 (severe insomnia). A cut-off score of 7 was used to determine relatively mild insomnia (Morin, 1993). The cut-off score was set at this level instead of the clinically more relevant cut-off score of 10 (Morin, Belleville, Bélanger, & Ivers, 2011) because the goal of this study was to demonstrate the feasibility of the app in a sample of patients with at least subclinical levels of insomnia.

Secondary measures

Pittsburgh Sleep Quality Index

The Pittsburgh Sleep Quality Index (PSQI) is a self-rating scale which measures sleep disturbances over a one-month period (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989). The PSQI consists of nineteen items with scores from 0 to 3 evaluating seven sub domains. The scores of these sub-domains are summed to calculate a global score ranging from 0 to 21. A global score above 5 indicates that someone is suffering severely in at least two domains, or moderately in at least three domains.

Dysfunctional beliefs about sleep

Dysfunctional beliefs were measured with the Dutch translation of the brief Dysfunctional Belief and Attitudes about Sleep scale (DBAS-16) (Morin, Vallières, & Ivers, 2007). The DBAS consists of sixteen statements with scores from 0 to 10 to indicate how much people agree with the statement. The average is calculated so that the total score ranges from 0 (no dysfunctional beliefs) to 10 (severe dysfunctional beliefs).

Anxiety Symptoms

Anxiety symptoms were assessed with the seven anxiety items of the Dutch version of the Hospital Anxiety and Depression Scale (HADS) (Spinhoven et al., 1997; Zigmond & Snaith, 1983). The summed score ranges from 0 (no symptoms of anxiety) to 21 (severe symptoms of anxiety).

Depression symptoms

Depressive symptoms were measured using a Dutch translation of the Centre of Epidemiological Studies Depression (CES-D) scale. The CES-D consists of twenty items with scores ranging from 0 to 3, which are summed, with higher scores indicating more depressive symptoms (Bouma, Ranchor, Sanderman, & van Sonderen, 1995; Radloff, 1977).

5.2.3.2. Diary measures

An online Dutch translated version of the consensus sleep diary (Carney et al., 2012) was used. Participants filled out the sleep diary for seven days. In the diary they recorded the time they went to bed, the time they tried to go to sleep, their time of final awakening, time out of bed, sleep onset latency (SOL), wake after sleep onset (WASO), terminal wakefulness (TWAK), number of awakenings (NWAK), sleep quality (SQ) (1 = “very bad” to 10 = “very good”), and use of sleep medication. From these variables, the time in bed (TIB = final arising time – time of going to bed), sleep time (TST = TIB - SOL - WASO - TWAK), and sleep efficiency (SE = (TST/TIB) × 100) were calculated. Sleep efficiency was the second primary measure in this trial.

5.2.3.3. Process measures

Furthermore, motivation to use the app and the acceptance of the app was measured by the Situational Motivation Scale (SIMS) (Guay, Vallerand, & Blanchard, 2000) and the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh, Morris, Davis, & Davis, 2003) respectively. The focus of this paper is on the outcome measures, so the results of the process measures are not included in this paper.

5.2.4. Procedure

Participants gave online informed consent and filled in the questionnaire containing the inclusion and exclusion criteria, demographic information, and the outcome measures. The participants who met the study criteria received a link to an online sleep diary by email for seven successive days. Emails for the diary were sent at 6.00 a.m. and a reminder email at 10.00 a.m. Participants who reported an average total sleep time of less than five hours were excluded, and the others were randomised to either the app or waiting list conditions. Randomisation was carried out by a third party who was not part of this study. They used an online tool[1] to generate blocks of twenty participants. The list of the randomisation sequence was kept in a locked office cupboard by the third party. After participants were assigned to a condition, participants and the principal investigator (CH) were no longer blind to the condition allocation.

Three weeks after starting with the app or the waiting-list condition, all participants received an interim measurement consisting of the ISI and DBAS supplemented with questions regarding motivation (SIMS) and app acceptance (UTAUT) for the app group. These interim measures are not reported in this paper. Seven weeks after randomisation both groups received a post-intervention questionnaire consisting of all the outcome measures (ISI, PSQI, DBAS, CES-D, HADS) and a seven-day diary. In addition, participants in the app group received questions regarding the effect and utility of the app, which are not reported in this paper. After completing the diary, participants in the waiting list condition received the app. Participants in the app condition additionally received a three-month follow-up questionnaire and diary. The study was approved by the internal Ethical Review Board of the University of Amsterdam, and was registered at trialregister.nl (NTR5560).

5.2.5. Statistical analysis

5.2.5.1. Required statistical power

To our knowledge this study is the first large-scale randomised controlled trial in which an app to treat chronic insomnia has been studied, so expected effects were unknown. Earlier research about online CBT-I found Cohen's *d* greater than 1.0 (Espie, Kyle, et al., 2012b; Lancee, Van Straten, Morina, Kaldo, & Kamphuis, 2015). It was uncertain

[1] http://www.randomization.com/

whether these large effects could also be obtained by an app, so an average effect was anticipated. A power calculation for a mixed ANOVA design (effect: f^2 = 0.15, power = 0.80, alpha-level = 0.05) indicated that a total of ninety participants were needed to detect a potential difference between the two conditions. As a meta-analysis showed that on average 50% of people adhere to technology-mediated insomnia treatment (Horsch, Lancee, Beun, Neerincx, & Brinkman, 2015), the goal was set to include 180 participants.

5.2.5.2. Analyses

The effects of the intervention were tested using multilevel analyses, which allows the inclusion of participants with one measurement and therefore is appropriate for intention-to-treat analyses (Hox, 2010). Models were built in R version 3.1.3 to explore within-group (time), between-group (condition: app vs. wl), and interaction (time x condition) effects. Model 0 was the basic model and included only the participants as a random intercept. Model 1 added the fixed factor time to model 0. Model 2 was built on model 1 and added the condition as a fixed effect. Finally, model 3 added the interaction effect between time and condition. Models 4 and 5 concern the pre-measurement and follow-up data. Model 4 is the null model that includes only the participants as a random intercept. In model 5 time is added. Since there were no follow-up measurements for the waiting list, condition is not included. Drop-out analyses for the post-measurements showed that age, SQ, and TWAK were associated with non-response in the app condition. In the waiting list condition NWAK was related with non-response. Drop-out analyses for the follow-up measurements showed that TWAK was associated with non-response. Therefore they were added as covariates in all models in the multilevel regression analyses (Hox, 2010). χ2s were calculated for the different models to compare the ability of the models to fit the data. Furthermore, Rs were calculated to indicate the level of variance explained by the level 1 variables (Snijders & Bosker, 1994). R-values of 0.10 indicate a small effect, r = 0.30 indicates a medium effect, and r = 0.50 indicates a large effect (Cohen, 1988).

To enhance comparability with other studies, between-group Cohen's ds were calculated. Table 2 shows the means and effect sizes based on an imputed dataset. First, within SPSS version 22 multiple imputation was used to insert missing cases (Sterne et al., 2009). Data from forty-one to forty-four participants (27.2–29.1%) was missing for the outcome questionnaires. Diaries were missing from seventy-six participants (50.3%). The follow-up measurements were not imputed due to a large amount of missing data. For imputation, the pre- and post-measures of the ISI, PSQI, DBAS, CES-D, HADS, SQ, SOL, WASO, NWAK, TIB, TWAK, TST, and SE were used, next to gender and age. With a predictive mean matching procedure ten separate datasets were generated. The values in Table 2 are based on these imputed datasets. Second, Cohen's *d*s were calculated by dividing the difference in change scores by the pooled standard deviation of that change score ($d = (M_{\text{change score waiting-list}} - M_{\text{change score app}})/\sigma_{\text{pooled}}$). Within-group Cohen's *d*s were calculated using the pre- and post-scores per condition and the pooled standard

deviation (d = (M_{pre} – M_{post})/σ_{pooled}). Additionally, within-group Cohen's ds were also calculated with the pre- and follow-up scores per condition and the pooled standard deviation (d = (M_{pre} – $M_{follow\text{-}up}$)/σ_{pooled}) (Table 3). A Cohen's d of 0.20 indicates a small effect, 0.50 a moderate effect, and 0.80 a large effect (Cohen, 1988).

5.3. Results

5.3.1. Baseline characteristics of the sample

Table 1 shows the baseline characteristics and the randomisation check. At baseline, the groups did not differ significantly on any demographic characteristics (all ps > .05).

5.3.2. Efficacy: Intention-to-treat analyses

Table 2 displays the mean scores for all the outcome measures and corresponding Cohen's ds for the baseline and post-measurements. Table 3 displays the mean scores for the follow-up measures. Figure 4 graphically depicts the scores for the main outcome measures ISI and SE. The results of the multilevel analyses are presented in Tables 4 and 5. Multilevel analyses showed significant interaction effects between time and condition on the primary outcome measures ISI (d = -.66) and sleep efficiency (d = .71) at post-test. These effects indicate that the app was more effective than the waiting list condition. Furthermore, WASO, NWAK, PSQI, CES-D, and HADS improved and showed significant interaction effects (Tables 4 and 5), but SOL, TIB, TWAK, TST, and DBAS showed no significant effects at post-test. At follow-up improvements on all outcome measures remained significant, except for NWAK.

5.3.2.1. Clinical changes

From the participants who completed the pre- and post-test, a clinically meaningful change was found on the Insomnia Severity Index (Δ ISI ≥ 8) (Morin et al., 2011) between the waiting list and the app conditions. A significant clinically meaningful change was observed twenty times in the app condition (20/45 = 44%) and seven times in the waiting list condition (7/62 = 11%) at the post-test. In the app condition significantly more people reached a meaningful clinical change (χ^2 (1) = 15.19, p < .001). Before treatment, all participants had an ISI score greater than 7 (Morin, 1993). Of the participants that completed the post-test, eighteen in the app condition (17/45 = 38%) and six in the waiting list condition (6/62 = 10%) had an ISI score less than or equal to 7. In the app condition significantly more participants dropped below the insomnia threshold of ISI less than or equal 7 than in the waiting-list condition (χ^2 (1) = 12.20, p < .001). At follow-up seven of the 29 participants had an ISI-score ≤ 7 (7/29 = 24%).

Table 2. Observed baseline and imputed post-test means and effect sizes

	Group	Baseline Mean (*SD*)	Post-test Mean (*SD*)	Cohen's *d* Within Group	Cohen's *d* Between group[a]
Questionnaire					
Insomnia severity	WL	16.4 (3.3)	13.2 (4.5)	.80	-.66
	App	16.4 (3.1)	9.9 (4.9)	1.33	
Dysfunctional beliefs	WL	5.2 (1.3)	4.8 (1.6)	.24	-.15
	App	5.3 (1.3)	4.7 (1.4)	.41	
Depression	WL	15.0 (5.8)	15.5 (9.5)	-.06	-.94
	App	16.5 (6.0)	11.0 (5.6)	.98	
Anxiety	WL	5.6 (3.1)	6.2 (3.8)	-.15	-.75
	App	6.1 (3.0)	4.1 (2.5)	.81	
PSQI	WL	10.6 (2.8)	9.7 (2.9)	.32	-.77
	App	11.0 (2.8)	7.4 (3.3)	1.09	
Diary					
Sleep efficiency	WL	77.0 (8.2)	78.3 (7.6)	-.16	.71
	App	77.6 (7.3)	84.8 (5.3)	-1.37	
Time in bed	WL	500 (46)	513 (34)	-.32	-.55
	App	506 (44)	495 (31)	.36	
Total sleep time	WL	386 (49)	401 (47)	-.32	.24
	App	393 (52)	421 (37)	-.74	
Sleep onset	WL	31 (21)	30 (19)	.06	-.45
	App	33 (20)	20(12.)	1.01	
Wake after sleep onset	WL	44 (30)	44 (25)	-.02	-.70
	App	45 (32)	24 (11)	1.79	
Terminal wakefulness	WL	37 (20)	37 (15)	.02	-.26
	App	35 (22)	29 (13)	.52	
Number of awakenings	WL	2.22 (1.14)	2.14 (1.08)	.07	-.38
	App	1.94 (0.99)	1.46 (1.14)	.42	
Sleep quality	WL	2.93 (0.52)	3.10 (0.55)	-.33	.29
	App	2.97 (0.41)	3.33 (0.54)	-.67	

[a] change scores

Table 3. Completers sample: baseline, post, and follow-up means for the app condition and within group effect sizes for the baseline and follow-up measurements

	Group	Baseline Mean (*SD*)	Post-test Mean (*SD*)	Follow-up Mean (*SD*)	Cohen's *d* Within Group Pre-Follow-Up
Questionnaire					
Insomnia severity	App	16.4 (3.1)	9.8 (4.8)	10.0 (5.3)	1.20
Dysfunctional beliefs	App	5.3 (1.3)	4.7 (1.4)	4.3 (1.8)	.58
Depression	App	16.5 (6.0)	10.3 (5.3)	11.0 (7.2)	.75
Anxiety	App	6.1 (3.0)	4.0 (2.4)	4.3 (2.8)	.67
PSQI	App	11.0 (2.8)	7.6 (3.1)	9.1 (3.6)	.53
Diary					
Sleep efficiency	App	77.6 (7.3)	83.8 (8.3)	83.8 (10.9)	-.57
Time in bed	App	506 (44)	496 (50)	483 (39)	.57
Total sleep time	App	393 (52)	417 (62)	403 (57)	-.17
Sleep onset	App	33 (20)	22 (14)	21 (15)	.80
Wake after sleep onset	App	45 (32)	27 (21)	25 (24)	.84
Terminal wakefulness	App	35 (22)	31 (22)	35 (36)	.13
Number of awakenings	App	1.94 (0.99)	1.58 (1.10)	1.75 (1.44)	.01
Sleep quality	App	2.97 (0.41)	3.38 (0.51)	3.41 (0.60)	-.74

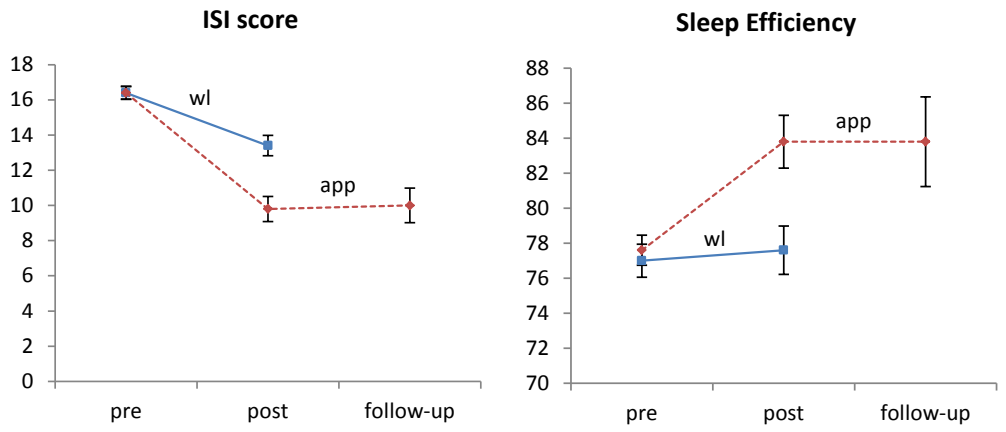


Figure 4. Completers sample: Insomnia Severity Index scores and Sleep Efficiency between groups on base-line, post-test, and three month follow-up. Error bars represent Standard Error.

Table 4. Multilevel analyses results of the diary variables

	SE		SOL		WASO		NWAK	
Model 3	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	77.38**	0.78	32.75**	2.47	42.87**	2.99	2.20**	0.13
Time	1.97*	0.99	-3.59	2.63	-3.74	3.79	-0.17	0.14
Condition	-0.24	1.11	0.12	3.52	3.17	4.27	-0.11	0.19
Interaction	5.34**	1.55	-7.86	4.16	-22.42**	5.97	-0.18*	0.22
	$\chi^2(1)$	*R*	$\chi^2(1)$	*R*	$\chi^2(1)$	*R*	$\chi^2(1)$	*R*
Model 0 vs 1	23.23**	.26	9.83**	.14	14.82**	.22	9.28**	.00
Model 1 vs 2	1.53	.10	0.39	.00	0.62	.00	1.89	.10
Model 2 vs 3	11.54**	.17	3.57	.10	13.29**	.17	4.63*	.10
Model 0 vs 3	36.30**	.33	13.79**	.17	28.73**	.28	15.81**	.14
Model 4 vs 5	18.66**	.21	12.21**	.12	11.99**	.14	3.59	.06

	TIB		TWAK		TST	
Model 3	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	499.74**	5.02	36.93**	2.35	387.6**	5.53
Time	9.75	5.22	-0.44	3.24	18.6**	6.18
Condition	5.46	7.16	-1.50	3.35	2.76	7.89
Interaction	-16.07	8.28	-3.23	5.08	14.9	9.77
	$\chi^2(1)$	*R*	$\chi^2(1)$	*R*	$\chi^2(1)$	*R*
Model 0 vs 1	0.69	.00	0.41	.00	23.45**	.26
Model 1 vs 2	0.03	.00	0.65	.10	0.85	.00
Model 2 vs 3	3.82	.10	0.41	.00	2.35	.00
Model 0 vs 3	4.54	.14	1.46	.10	26.66**	.28
Model 4 vs 5	5.38*	.10	0.85	.00	4.35*	.05

*p <.05, **p < .01

NWAK = number of awakenings, SE = Sleep efficiency, SOL = Sleep Onset Latency, TIB = Time In Bed, TST = Total Sleep Time, TWAK = Terminal wakefulness, WASO = Wake After Sleep Onset

The covariates are not reported in this table.

Table 5. Multilevel analyses results of the questionnaire variables

	ISI		PSQI		DBAS	
Model 3	*B*	*(SE)*	*B*	*(SE)*	*B*	*(SE)*
Intercept	16.26**	0.44	10.50**	0.33	5.17**	0.16
Time	-3.04**	0.53	-0.97**	0.36	-0.34*	0.15
Condition	0.20	0.63	0.56	0.47	0.18	0.22
Interaction	-3.53**	0.79	-2.52**	0.54	-0.31	0.23
	$\chi^2(1)$	*R*	$\chi^2(1)$	*R*	$\chi^2(1)$	*R*
Model 0 vs 1	86.00**	.50	43.05**	.33	16.32**	.14
Model 1 vs 2	4.65*	.14	0.91	.10	0.12	.00
Model 2 vs 3	19.07**	.22	20.65**	.22	1.86	.00
Model 0 vs 3	109.73**	.56	64.61**	.39	18.31**	.14
Model 4 vs 5	66.02**	.55	9.28**	.21	11.65**	.22

	CES-D		HADS	
Model 3	*B*	*(SE)*	*B*	*(SE)*
Intercept	14.95**	0.79	5.69**	0.35
Time	0.66	0.93	0.41	0.43
Condition	1.54	1.12	0.38	0.50
Interaction	-6.58**	1.40	-2.36**	0.65
	$\chi^2(1)$	*R*	$\chi^2(1)$	*R*
Model 0 vs 1	8.34**	.14	3.17	.10
Model 1 vs 2	0.96	.10	1.49	.10
Model 2 vs 3	20.67**	.22	12.75**	.20
Model 0 vs 3	29.97**	.28	17.41**	.24
Model 4 vs 5	18.66**	.28	10.03**	.17

*p <.05, **p < .01
CES-D = Centre of Epidemiological Studies Depression scale, DBAS = Dysfunctional Beliefs About Sleep, HADS = Hospital Anxiety and Depression Scale , ISI = Insomnia Severity Index, PSQI = Pittsburgh Sleep Quality Index
The covariates are not reported in this table.

5.3.3. Treatment adherence

Treatment adherence was divided in four components: 1) the number of sleep diaries filled out; 2) the number of relaxation exercises performed; 3) the deviation between real time in bed and agreed-upon time in bed; and 4) the number of conversations completed (Table 6). Two of the seventy-four participants failed to download the app. Furthermore, the adherence data showed different adherence patterns (see Appendix M). Most participants (n = 35) filled in more than thirty-five diaries, thirteen participants filled in less than seven diaries, and the other twenty-four participants filled in between seven and thirty-five diaries. This pattern follows a U-shape. For the relaxation exercise another pattern can be distinguished. More than half of the participants (n = 41) performed a maximum of seven relaxation exercises, of which eleven participants did not do any relaxation exercises at all. Seven participants did more than thirty-five relaxation exercises. For the conversations the opposite was true; most participants (n = 47) finished 90% or more of the conversations. Only a few participants (n = 4) finished less than 10% of the conversations. Thirty-eight participants started and committed to the sleep restriction exercise, meaning that they came to an agreement with the coach about their time in bed. Participants could stay in bed for longer or for less time than the agreement, and both situations occurred. Thirty-two of the thirty-eight participants stayed in bed for longer than agreed on for most of the nights involving sleep restriction. Where participants stayed in bed too long it was by 67 minutes on average (SD = 45); when they shortened their time in bed it was by 42 minutes on average (SD = 29). Of the thirty-eight participants twenty-six (68.4%) were adherent, meaning that their time in bed deviated by an average of less than 60 minutes from the agreed-upon time in bed.

5.4. Discussion and conclusion

In this large-scale randomised controlled trial we investigated the efficacy of CBT-I delivered via a smartphone app. The results show that the app had moderate significant effects compared to a waiting list on the primary measures of insomnia severity (d = -0.66) and sleep efficiency (d = 0.71). On the secondary measures the following measures also improved compared to the waiting list: wake after sleep onset, number of awakenings, Pittsburgh Sleep Quality Index, depression and anxiety decrease. At post-test 44% of the participants in the app condition had achieved a clinically meaningful change compared to 11% in the waiting list condition. The improvements were largely sustained at three-month follow-up. The observed effects on the primary measures are similar to those reported in a very recent meta-analysis on sleep efficiency (Hedge's g = 0.58) and somewhat lower (but in the same range) than those reported for insomnia severity (Hedge's g = 1.09) (Zachariae et al., 2016). Note that the meta-analysis was based on studies with various levels of human involvement, ranging from no human support to personal contact as part of the intervention. Earlier research indicated that human support increases efficacy (Lancee, van den Bout, Sorbi, & van Straten, 2013). However, the effect sizes in the current study were achieved without any form of human support.

Table 6. Treatment adherence

Treatment adherence	*M (SD)*	range	# of ss with adequate dose*
People who downloaded the app	72 of the 74 participants		
Number of diaries filled in (n=72)	29.1 times(16.4) /49 = 59.4%	0-48 times	35/72 = 48.6%
Number of relaxation exercises performed (n=72)	10.8 times (12.0) /49 = 22.0%	0-45 times	7/72 = 9.7%
Completed conversations of the training (n=72)	83% (27%)	0-100%	47/72 = 65.3%
Deviation from sleep restriction in minutes (n=38)	59.2 minutes (46.4)	9-285 min	26/38 = 68.4%

* adequate dose: diaries > 35, relaxation exercises > 35, deviation sleep restriction < 60 minutes, conversations > 90%

Regarding automated support, this study most closely resembles the trials by Espie and colleagues (2012a) and Ritterband and colleagues (2009) which both offered automated online CBT-I. These online treatments packaged the full scope of CBT-I and demonstrated large effects. Espie and colleagues found a Cohen's d of 0.95 for sleep efficiency. Ritterband and colleagues found Cohen's ds of 1.26 for insomnia severity and 0.68 for sleep efficiency. Again, the observed effect sizes in the current study were more or less in the same range as these published results, and our effects were achieved without including the full CBT-I package (e.g. cognitive therapy and stimulus control were not included). The app concentrated on sleep restriction, and as a result the effects for sleep efficiency are more pronounced than those for insomnia severity. The focus on sleep restriction may also explain the absence of an effect on TST.

Zachariae and colleagues (Zachariae et al., 2016) found in their meta-analysis that 58.7% to 100% of the participants in the CCBT-I conditions completed post-intervention assessments, with an average of 75.3%. In the current study 60.8% of the participants in the app condition completed post-intervention assessment questionnaires, while 80.6% of the participants in the waiting list condition did so. This difference can probably be explained by the fact that the participants in the waiting list only received the app after they had filled in the post-intervention assessment. However, the number of participants that fills in assessments may not necessarily correspond to the number of participants that completes the interventions. Therefore, treatment adherence numbers and adequate doses are also reported. Previously, Espie and colleagues (2012a) found that 88% of their participants received an adequate dose (≥ 4 sessions). Lancee and colleagues (2013) found that 83% received an adequate dose of the modules in the support condition, and 60% in the no support condition. In this trial adherence

was measured for the different components, with adherence rates fluctuating between 9.7% and 68.4%. Apart from the relaxation exercise adherence (where only 9.7% of the participants received an adequate dose), the other adherence rates are comparable to the 60% found by Lancee and colleagues (2013) in their no support condition. In general adherence rates were adequate, but there was also a considerable number of people who did not start the modified sleep restriction exercise at all. Beforehand, the decision was made that it was better to keep people in no or a suboptimal sleep restricting schedule rather than letting them drop out of the treatment altogether. However, the optimal tradeoff between individual autonomy and strictness in smartphone app regimes has yet to be determined in future studies.

5.4.1. Limitations and future work

This study has a number of limitations that should be considered in relation to the findings. Since the goal of the study was to demonstrate the efficacy of the app first in a group with insomnia disorder but without too much sleep impairment, we used an ISI score of > 7, meaning that people who slept fewer than five hours as measured by a sleep diary were excluded. This exclusion criterion may have led to a floor effect and the inclusion of participants with relatively little room for improvement. Although it is hard to compare the different studies because of different inclusion criteria, it seems that Espie and colleagues (Espie, Kyle, et al., 2012a) only included participants with more severe insomnia (baseline sleep efficiency of 55%–65%). It is arguably possible to achieve larger effects in samples with higher levels of symptoms. However, it remains the case that the efficacy of our smartphone app has not yet been demonstrated in a sample with severe insomnia. Because this was one of the first times a standalone app has been used to deliver CBT-I, participants with comorbidities such as depression were also excluded. This and the issues mentioned above limit the generalisability of our results, especially given the high comorbidity of depression and insomnia. Now that the app has proven its efficacy in a relatively mildly affected sample, future research could expand the inclusion criteria (e.g. severe insomnia, depression) to study the effectiveness of a CBT-I app in a more severely affected population.

A methodological limitation was that no other online or face-to-face treatment group was included. Several other studies have already demonstrated the efficacy of CCBT-I and CBT-I programmes. However, a similar online condition could provide insight into the added value of a mobile app. Another related limitation is that there was no placebo control group. It may very well be that non-specific factors have played a role in the treatment effects of the app. Other methodological limitations were that this study used self-report measures, and polysomnography would be needed to confirm the objective changes in sleep. Furthermore, the participants in this study were a self-selected sample and may represent an unusual group of people that is interested in solving their sleep difficulties with self-help. The results show that a high percentage of the sample consists of university educated participants, which only represents a part of society.

Another limitation was that the app focused on sleep restriction and relaxation. Future work should include more of the other CBT-I components, for example cognitive exercises, and evaluate those. Smartphone apps provide us with the unique opportunity to study the separate components of CBT-I in a controlled way. Future research could focus on studying the separate components, so more insight is gained regarding the individual effectiveness of these CBT-I components.

Lastly, there were some technical issues during the RCT which made it impossible for some participants to continue to the next conversation. The occurrence of this problem was monitored and solved when needed. In these cases a new conversation was manually planned in the database for a specific participant, and an email with instructions to update the app was sent to that participant.

5.4.2. Conclusion

We are confident that this study has produced insights in the domain of automated e-coaching apps for insomnia. These applications provide an opportunity to investigate separate treatment components while minimising the influence of non-specific therapist factors such as therapeutic alliance. Keeping the limitations in mind, this study demonstrated the efficacy of a smartphone app in the treatment of insomnia. These effects were clinically meaningful and in the range of what is found for online treatment in general. This supports the applicability of these kinds of technical tools in the treatment of insomnia. Through these apps, many more people can be offered effective insomnia treatment with probable reduced costs. We are confident that smartphone apps will prove to be useful in the realm of prevention treatments; it remains to be determined how they should best be offered, e.g. in a standalone format for (prevention) treatment, or within a blended care framework where the sleep specialist uses an app to improve and accelerate insomnia treatment.

5.5. References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed ed.). Arlington, VA: American Psychiatric Publishing.

Baglioni, C., Battagliese, G., Feige, B., Spiegelhalder, K., Nissen, C., Voderholzer, U., . . . Riemann, D. (2011). Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. J Affect Disord, 135(1), 10-19.

Bauer, J., Consolvo, S., Greenstein, B., Schooler, J., Wu, E., Watson, N. F., & Kientz, J. (2012). ShutEye: encouraging awareness of healthy sleep recommendations with a mobile, peripheral display. Paper presented at the CHI 2012, Austin, Texas, USA.

Beun, R. J., Brinkman, W. P., Fitrianie, S., Griffioen-Both, F., Horsch, C., Lancee, J., & Spruit, A. G. L. (Accepted 2016). Improving adherence in Automated e-Coaching. A Case from Insomnia Therapy. Paper presented at the Persuasive 2016, Salzburg, Austria.

Beun, R. J., Griffioen-Both, F., Ahn, R., Fitrianie, S., & Lancee, J. (2014). Modeling Interaction in Automated E-Coaching-A Case from Insomnia Therapy. Paper presented at the COGNITIVE 2014, The Sixth International Conference on Advanced Cognitive Technologies and Applications.

Bouma, J., Ranchor, A. V., Sanderman, R., & van Sonderen, E. (1995). Het meten van symptomen van depressie met de CES-D: Een handleiding [Dutch translation of the Epidemiological Studies-Depression scale]. Groningen: Noordelijk Centrum voor Gezondheidsvraagstukken.

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry research, 28(2), 193-213.

Carney, C. E., Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Krystal, A. D., Lichstein, K. L., & Morin, C. M. (2012). The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep, 35(2), 287.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. 2nd edn. Hillsdale, New Jersey: L: Erlbaum.

Consolvo, S., McDonald, D. W., & Landay, J. A. (2009). Theory-driven design strategies for technologies that support behavior change in everyday life. Paper presented at the CHI 2009, Boston, Massachusetts, USA.

Daley, M., Morin, C. M., LeBlanc, M., Grégoire, J.-P., & Savard, J. (2009). The economic burden of insomnia: direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms, and good sleepers. Sleep, 32(1), 55.

Eijk, R. M. v. (2013). Ambient Coaching of Progressive Relaxation. Paper presented at the Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 03.

Espie, C., Kyle, S., Williams, C., Ong, J., Douglas, N., Hames, P., & Brown, J. (2012a). A randomized, placebo-controlled, trial of online Cognitive Behavioral Therapy for chronic Insomnia Disorder delivered via an automated media-rich web application. Sleep, 35(6), 769.

Espie, C. A., Kyle, S. D., Hames, P., Cyhlarova, E., & Benzeval, M. (2012). The Daytime Impact of DSM-5 Insomnia Disorder: Comparative Analysis of Insomnia Subtypes From the Great British Sleep Survey. The Journal of clinical psychiatry, 73(12), 1,478-1484.

Espie, C. A., Kyle, S. D., Williams, C., Ong, J. C., Douglas, N. J., Hames, P., & Brown, J. S. L. (2012b). A randomized, placebo-controlled trial of online cognitive behavioral therapy for chronic insomnia disorder delivered via an automated media-rich web application. Sleep, 35(6), 769-781. doi: 10.5665/sleep.1872

Griffioen-Both, F., Beun, R., Fitrianie, S., Spruit, A. G. L., Horsch, C. H. G., & Lancee, J. ((in preparation)). Shared decision-making in automated e-coaching.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). Motivation and emotion, 24(3), 175-213.

Harvey, L., Inglis, S. J., & Espie, C. A. (2002). Insomniacs' reported use of CBT components and relationship to long-term clinical outcome. Behaviour research and therapy, 40(1), 75-83.

Ho, F. Y.-Y., Chung, K.-F., Yeung, W.-F., Ng, T. H., Kwan, K.-S., Yung, K.-P., & Cheng, S. K. (2014). Self-help cognitive-behavioral therapy for insomnia: A meta-analysis of randomized controlled trials. Sleep Medicine Reviews.

Horsch, C. H. G., Lancee, J., Beun, R. J., Neerincx, M. A., & Brinkman, W. P. (2015). Adherence to Technology-mediated Insomnia Treatment: A meta-analysis, interviews with users, and focus groups with users and experts. Journal of Medical Internet Research, 17(9), e214. doi: 10.2196/jmir.4115

Hox, J. J. (2010). Multilevel Analysis, Techniques and Applications (2nd ed.). New York: Routledge.

Irwin, M. R., Cole, J. C., & Nicassio, P. M. (2006). Comparative meta-analysis of behavioral interventions for insomnia and their efficacy in middle-aged adults and in older adults 55+ years of age. Health Psychol, 25(1), 3-14. doi: 10.1037/0278-6133.25.1.3

Kaltenthaler, E., & Cavanagh, K. (2010). Computerised cognitive behavioural therapy and its uses. Progress in Neurology and Psychiatry, 14(3), 22-29.

Konrath, S., & Yan, Z. (2015). Positive technology. Using mobile phones for psychosocial interventions. Encyclopedia of Mobile Phone Behavior.

Kyle, S. D., Aquino, M. R. J., Miller, C. B., Henry, A. L., Crawford, M. R., Espie, C. A., & Spielman, A. J. (2015). Towards standardisation and improved understanding of sleep restriction therapy for insomnia disorder: A systematic examination of CBT-I trial content. Sleep Medicine Reviews, 23, 83-88.

Kyle, S. D., Morgan, K., & Espie, C. A. (2010). Insomnia and health-related quality of life. Sleep Medicine Reviews, 14(1), 69-82.

Lancee, J., van den Bout, J., Sorbi, M. J., & van Straten, A. (2013). Motivational support provided via email improves the effectiveness of internet-delivered self-help treatment for insomnia: A randomized trial. Behaviour research and therapy, 51(12), 797-805.

Lancee, J., Van Straten, A., Morina, N., Kaldo, V., & Kamphuis, J. (2015). Guided Online or Face-to-Face Cognitive Behavioral Treatment for Insomnia? A Randomized Wait-list Controlled Trial. Sleep.

Lawson, S., Jamison-Powell, S., Garbett, A., Linehan, C., Kucharczyk, E., Verbaan, S., . . . Morgan, K. (2013). Validating a mobile phone application for the everyday, unobtrusive, objective measurement of sleep. Paper presented at the Conference on Human Factors in Computing Systems - Proceedings.

LeBlanc, M., Beaulieu-Bonneau, S., Mérette, C., Savard, J., Ivers, H., & Morin, C. M. (2007). Psychological and health-related quality of life factors associated with insomnia in a population-based sample. Journal of psychosomatic research, 63(2), 157-166.

Mallon, L., Broman, J.-E., & Hetta, J. (2005). High Incidence of Diabetes in Men With Sleep Complaints or Short Sleep Duration A 12-year follow-up study of a middle-aged population. Diabetes Care, 28(11), 2762-2767.

Miller, C. B., Espie, C. A., Epstein, D. R., Friedman, L., Morin, C. M., Pigeon, W. R., . . . Kyle, S. D. (2014). The evidence base of sleep restriction therapy for treating insomnia disorder. Sleep Medicine Reviews, 18(5), 415-424.

Morin, C. M. (1993). Insomnia: Psychological assessment and management: Guilford Press New York.

Morin, C. M., Belleville, G., Bélanger, L., & Ivers, H. (2011). The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep, 34(5), 601.

Morin, C. M., Bootzin, R. R., Buysse, D. J., Edinger, J. D., Espie, C. A., & Lichstein, K. L. (2006). Psychological and behavioral treatment of insomnia: update of the recent evidence (1998-2004). SLEEP-NEW YORK THEN WESTCHESTER-, 29(11), 1398.

Morin, C. M., & Espie, C. A. (2003). Insomnia: A clinician's guide to assessment and treatment (Vol. 1): Springer.

Morin, C. M., Hauri, P. J., Espie, C. A., Spielman, A. J., Buysse, D. J., & Bootzin, R. R. (1999). Nonpharmacologic treatment of chronic insomnia. An American Academy of Sleep Medicine review. Sleep, 22(8), 1134-1156.

Morin, C. M., Vallières, A., & Ivers, H. (2007). Dysfunctional beliefs and attitudes about sleep (DBAS): validation of a brief version (DBAS-16). Sleep, 30(11), 1547.

Ohayon, M. M. (2002). Epidemiology of insomnia: what we know and what we still need to learn. Sleep medicine reviews, 6(2), 97-111.

Radloff, L. S. (1977). The CES-D scale a self-report depression scale for research in the general population. Applied psychological measurement, 1(3), 385-401.

Ritterband, L. M., Thorndike, F. P., Gonder-Frederick, L. A., Magee, J. C., Bailey, E. T., Saylor, D. K., & Morin, C. M. (2009). Efficacy of an Internet-based behavioral intervention for adults with insomnia. Archives of General Psychiatry, 66(7), 692.

Shirazi, A. S., Clawson, J., Hassanpour, Y., Tourian, M. J., Schmidt, A., Chi, E. H., . . . Van Laerhoven, K. (2013). Already Up? Using Mobile Phones to Track & Share Sleep Behavior. International Journal of Human-Computer Studies.

Snijders, T. A., & Bosker, R. J. (1994). Modeled variance in two-level models. Sociological methods & research, 22(3), 342-363.

Spinhoven, P., Ormel, J., Sloekers, P., Kempen, G., Speckens, A., & Hemert, A. v. (1997). A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. Psychological medicine, 27(02), 363-370.

Spoormaker, V. I., Verbeek, I., van den Bout, J., & Klip, E. C. (2005). Initial validation of the SLEEP-50 questionnaire. Behavioral Sleep Medicine, 3(4), 227-246.

Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., . . . Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. Bmj, 338, b2393.

Suka, M., Yoshida, K., & Sugimori, H. (2003). Persistent insomnia is a predictor of hypertension in Japanese male workers. Journal of occupational health, 45(6), 344-350.

Taylor, D. J., Lichstein, K. L., Durrence, H. H., Reidel, B. W., & Bush, A. J. (2005). Epidemiology of insomnia, depression, and anxiety. SLEEP-NEW YORK THEN WESTCHESTER-, 28(11), 1457.

van Straten, A., & Cuijpers, P. (2009). Self-help therapy for insomnia: a meta-analysis. Sleep Medicine Reviews, 13(1), 61-71.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.

Verbeek, I., & Klip, E. C. (2005). Slapeloosheid [Insomnia]. Amsterdam: Boom.

Voinescu, B. I., Szentagotai, A., & David, D. (2013). Internet-administered cognitive-behavioral therapy for insomnia. Journal of Cognitive and Behavioral Psychotherapies, 13(1 A), 225-237.

Zachariae, R., Lyby, M. S., Ritterband, L. M., & O'Toole, M. S. (2016). Efficacy of internet-delivered cognitive-behavioral therapy for insomnia–A systematic review and meta-analysis of randomized controlled trials. Sleep Medicine Reviews, 30, 1-10.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. Acta psychiatr scand, 67(6), 361-370.

6. Conclusion

6.1. Conclusion

This thesis investigated how a virtual coach can provide therapy support to people suffering from insomnia. A virtual coach delivered via a smartphone app that encompassed sleep restriction, sleep diaries, relaxation exercises, sleep hygiene, education, reminders and negotiation, was proposed. The research described in this thesis was designed to answer the main research question:

> In what way can persuasive strategies contribute to improve treatment adherence to, and consequently the effectiveness of, a CBT-I-based virtual sleep coach?

The concepts that were considered in this collection of studies were adherence, adherence enhancing mechanisms, and the effectiveness of the proposed virtual sleep coach. Four hypotheses regarding these concepts were formulated to answer the main question:

H1: There is a positive relationship between adherence to a technology-mediated sleep intervention and the effectiveness of that intervention.

H2: A self-reflection tool can help people progress through the readiness-to-change stages.

H3: Computer-generated reminders increase adherence rates in technology-mediated sleep interventions.

H4: A fully automated virtual sleep coach app, encompassing sleep restriction, sleep diary, relaxation exercises, sleep hygiene, education, reminders and negotiation, is clinically effective in reducing insomnia.

The results presented in this thesis provide an insight into what way persuasive strategies can enhance adherence (hypotheses 1-3). Furthermore, they demonstrate that the proposed virtual coach is effective in reducing insomnia (hypothesis 4). The main conclusions are structured by examining the arguments for these four hypotheses.

6.1.1. The relationship between adherence and effectiveness in technology-mediated sleep interventions.

The conducted meta-analysis showed a positive relationship between treatment adherence and effectiveness of technology-mediated sleep interventions supporting hypothesis one. Furthermore, the meta-analysis resulted in more insights and knowledge regarding adherence. First, the meta-analysis did not find a relation between

experimental compliance and treatment adherence. Therefore, the position is taken that more attention should be paid to separately measuring and reporting treatment adherence and experimental compliance. Secondly, the meta-analyses showed an average adherence rate of 52% for technology-mediated sleep interventions.

The results of the meta-analysis were complemented with the outcomes of 15 interviews and 6 focus groups in order to gain insight into what way persuasive strategies could contribute to treatment adherence. During the interviews it proved to be difficult for the interviewees to identify adherence-enhancing mechanisms. However, they were able to give reasons for adherence and non-adherence to the technology-mediated sleep products they had been using. Reasons for adherence were: need of the functionality and personal beliefs, attitudes, and willpower. Reasons mostly mentioned for non-adherence were: doubting the effectiveness, no need for the functionality, and forgetfulness. These reasons provided insights into how persuasive strategies could contribute to treatment adherence. For example, the virtual sleep coach could include a needed functionality like an alarm clock. Furthermore, the interviewees articulated they were adhering quit well to the technology-mediated sleep products they had been using. Revealing a discrepancy between the found adherence rate in the meta-analysis and perceived adherence. Besides interviews, focus groups were organised to discuss scenarios that included the proposed virtual sleep coach app encompassing adherence enhancing strategies. The results showed that participants found the concepts “users in control” and “doing it for your own sake” reliable principles for adherence. The focus groups suggested that persuasive strategies could improve adherence if the found concepts are respected. Furthermore, the focus groups provided insights in the advantages and disadvantages of adherence enhancing strategies. For example, awarding points was not seen as a beneficial mechanism in sleep therapy by the participants. In summary, the meta-analysis indicated there is room for improving adherence rates to technology-mediated sleep interventions, while the qualitative studies suggested that participants did not consider adherence as a problem and they believe willpower to be an effective adherence strategy. This latter result should be interpreted with caution, because there seems to be an “adherence bias”. This thesis suggests that a virtual coach should be able to cope with this “adherence bias”, and persuade users to accept adherence-enhancing strategies.

6.1.2. A self-reflection tool can help people progress through the readiness-to-change stages .

The possibilities for a virtual coach to change someone’s readiness-to-change stage were explored in a field study. For this study a stage-matched self-reflection tool based on motivation interviewing and aimed at dealing with ambivalence around sleep was developed. Two versions of the tool were realized to explore whether a paper workbook or a digital workbook including a virtual coach could better support people progressing through the stages. Unfortunately, approximately half of the participants did use the workbook less than the prescribed two times. Qualitative data revealed that people are

not only ambivalent about their behaviour change, but also about the interventions and tools that support these change processes. A tool can be ambivalent for the user, i.e., can bring about variant perceived or experienced cost-benefit trade-offs. Tool ambivalence may be able to explain non-adherence. Current behaviour change theories, like the theory of planned behaviour (Ajzen, 1991), theory of reasoned action (Ajzen & Fishbein, 1988), social-cognitive theory (Bandura, 1977), cognitive dissonance theory (Festinger, 1957), the COM-B model (Michie, van Stralen, & West, 2011; Ripple, 1955), and the BMAT model (Fogg, 2009), focus on the behaviour itself. They explain how and when behaviour occurs, but none of them includes tool ambivalence. The results from this field study suggest that a virtual coach, even though it embeds persuasive strategies, can still fall prey to non-adherence because of tool ambivalence. Therefore, a virtual coach should take tool ambivalence into account, for example by reducing usage 'costs', increasing the perceived benefits, or tailor the intervention to personal drives.

6.1.3. Reminders increase adherence rates in technology-mediated sleep interventions.

The third hypothesis was supported by the results of another field experiment in which the underlying principles of two types of reminders were studied. An automatic reminder was based on the principles of the Capability-Opportunity-Motivation-Behaviour (COM-B) model, whereas a self-set reminder was based on several ideas regarding self-empowerment. Both reminders increased the sleep diary adherence. Additionally, mediation analyses showed that at least part of the effect could be explained by the underlying mechanisms of COM-B and self-empowerment. Meaning that an increase in perceived self-empowerment was associated with an increase in adherence, and a reminders given at better opportunities also increase adherence. In conclusion, the reminders can be an effective persuasive strategy to enhance adherence. COM-B and self-empowerment explain in what way reminders can enhance treatment adherence in an effective virtual sleep coach.

6.1.4. A virtual sleep coach is effective in reducing insomnia.

A randomized controlled trial was conducted to investigate the efficacy of a virtual sleep coach app. The results showed that the app had significant moderate effects on insomnia severity and sleep efficiency, which were the two main outcome measures. The majority of the other sleep measures also improved. The improvements were retained at the 3-month follow-up. These results demonstrate that a fully automated virtual sleep coach app, encompassing sleep restriction, sleep diary, relaxation exercises, sleep hygiene, education, reminders and negotiation, is clinically effective in reducing insomnia.

6.2. Limitations

To appreciate the work presented in this thesis, it is important to consider its limitations. First of all this research was conducted in the Netherlands with a self-selected sample in the domain of sleep. This abates the external validity of this research. Generally, people who were registered as participants at the Sleepcare project website were invited to participate in the studies. The Sleepcare website and the studies were promoted via social media, online advertisements, university lectures, flyers, posters, a press release, and personal connections. The people who participated in the studies probably deviated from the general population on a few points: a) they probably thought of themselves as having a sleep problem, b) they were pro-active and c) they were probably technology minded. In the RCT (chapter 5), for example, an unusual high percentage of the participants was university educated, which does not represent the general population. Furthermore, there were probably mainly tech-savvy people participating in the studies, since most recruitment was done online and the studies involved either technology-mediated sleep products, a computer program, or a mobile app. Although, earlier studies indicated that people are interested in mental health apps (Proudfoot et al., 2010; Torous, Friedman, & Keshavan, 2014), participating in this research required self-initiative (e.g., the app was not recommended by general practitioners). Therefore, the participants in the studies are likely to be innovators and early adopters (Rogers, 2010). Nevertheless, the characteristics of the self-selected sample most likely represent the people who would use a virtual sleep coach in a non-experimental set-up. In this sense, the sample is probably representative.

Furthermore the research was conducted in the Netherlands which might have influenced the level of acceptability of certain persuasive strategies. In general, Western culture tends to be individualistic, autonomous-focused, and directed at individual achievements (Varnum, Grossmann, Kitayama, & Nisbett, 2010). Therefore, for example, the concepts “users in control” and “doing it for your own sake” found in the focus groups are probably influenced by Dutch culture. If this research was performed in another country, different values might have been found.

Lastly, all the research in this thesis regarded sleep interventions. It might be the case that the results are not applicable to other health domains. Sleep is different from other health problems such as substance abuse, depression, or post-traumatic stress disorder. For instance, sleep is really intimate, people let down their guards and are defenceless when sleeping. Specific characteristics of the sleep domain might make that the results found in this research are not carried over to other health domains. For example, reasons for (none-)adherence might differ in other health domains. Other results, however, like the positive relationship between adherence and effectiveness probably also hold in other domains. In summary, conclusions regarding persuasive strategies and virtual coaches found in these thesis cannot be generalized to other health domains, other cultures, or other populations without some caution.

Another important limitation of this research is that the systems and interventions were investigated in their entirety. As a consequence, it is not possible to attribute specific results to particular components of the system or intervention. The system and intervention remain a so-called black box. One advantage, however, is that from these kinds of studies can be concluded that the system or intervention as a whole does work. Another advantage is that the studies could be done in the real world rather than in the laboratory, which increases the external validity. Related to this black box limitation is the limitation that only one app with certain design solutions was investigated. The app followed a 'Talk & Tools' design rationale (Fitrianie, Griffioen-Both, & Beun, in preparation). This means that a dialogue system supported most communication between the virtual coach and the user. While the CBT-I exercises were mainly accessible via the tools menu. The effects of the design choices were not investigated in this thesis. However, common mechanisms in Android apps were followed to the best of our abilities, in order to minimize the effects of the design on the outcome measures.

Lastly, this research heavily relies on questionnaires and sleep diaries. The disadvantages of questionnaires are that it is self-reported and with hindsight. However, some concepts cannot be measured otherwise, as by questionnaires. Per concept the best feasible method for measuring was chosen. For instance, insomnia is not defined by the hours of sleep someone gets, it is partly a subjective problem. The DSM criteria for insomnia are: a) disruptive sleep, for example having problems with falling asleep, staying asleep, or early wakefulness, b) the complaints need to be present for at least three nights a week for more than three months, and c) the complaints impair daily functioning (American Psychiatric Association, 2013). The latter criteria is a subjective measure, and therefore can best be measured by a questionnaire. And although polysomnography is the golden standard to measure sleep, sleep diaries have been widely used as a sleep measure as well (Buysse, Ancoli-Israel, Edinger, Lichstein, & Morin, 2006). Another example in which objective measures cannot be applied, is the exploration of reasons for (none-) adherence. It is hard to measure attitudes, beliefs, and reasons in a different way than just asking the participants. Moreover, not only subjective measures were used. For example, in the studies that included the app adherence was measured objectively by log-files on the smartphone. In conclusion, predominantly subjective measures were used in this research, the different measurement methods were tuned to the various concepts.

6.3. Contributions

6.3.1. Scientific contribution

The main contributions of this thesis are the insights gained regarding adherence, persuasive strategies and effectiveness of a virtual sleep coach. The thesis suggests a relationship between adherence and effectiveness, potential room to improve adherence rates, possible improvement supported by reminders, and that a virtual

coach could be effective. Also, this thesis brought new evidence to an old case. Most scholars state that adherence is important for effectiveness and that adherence can be improved, but no evidence was available in the domain of technology-mediated sleep interventions. The meta-analysis supports both these assumptions. The meta-analysis did not find a relationship between experimental compliance and treatment adherence. Therefore, it is important to distinguish these two measures. It is important that future studies measure and report treatment adherence rates, and not solely experimental compliance, since treatment adherence is related to effectiveness, and experimental compliance seems unrelated to treatment adherence. Furthermore, empirical work was done on the underlying principles of reminders that was not be done before. Self-empowerment and opportunity can partly explain why people follow up on reminders and perform the desired activity. This result provides us with deeper understanding of why reminders work, and might explain situations in which reminders do not work. In addition, a new concept of tool ambivalence was established. Tool ambivalence is not directly addressed in current behavior change theories, but similarities can be found in procrastination research, decision theories, and learning theories. Lastly, this work is original because it was the first study that investigated cognitive-behavioural-therapy for insomnia delivered through an app. Additionally, CBT-I was once more validated as an effective treatment for people with insomnia. Besides, the results of the RCT suggest that the individual components of CBT-I are effective on the outcome measures that is targeted by a specific component. For example, the app primarily focused on sleep restriction, and rather big differences in sleep efficiency were found. Furthermore, the app did not include any cognitive therapy, and no changes in dysfunctional beliefs about sleep (DBAS) were found. Generally, the scientific contributions were fetched by carrying out new empirical work, and by combining different research methods, such as quantitative studies and qualitative methods.

6.3.2. Practical contribution

The results of the studies and gained insights are not only relevant for the scientific community. Also therapist, patients, developers and designers of health apps can benefit from the results.

6.3.2.1. For therapists

First and foremost, the app demonstrated to be effective. This showcases that treatments can be offered through smartphones. With the current proliferation of health apps, it is hard to know for healthcare providers which apps work and which do not work. Firstly, the Sleepcare app is based on the treatment of recommendation for insomnia: CBT-I. Secondly, the app has been tested in an RCT with positive results, and therefore could be recommended to patients to support insomnia treatment.

6.3.2.2. For patients

Because the virtual coach was casted in a smartphone app, accessibility to insomnia treatment has improved. Instead of having to go to a therapist, the app is available to everyone possessing an Android smartphone. Furthermore, an app provides the opportunity to be flexible in use, independent of time and place. Of course, accessibility could even be further improved by developing the app also for other platforms and in other languages, but that was not part of the research goals.

6.3.2.3. For developers and designers

From every study done in this thesis a few insights for developing a virtual sleep coach were gained. In short, concepts that designers of a virtual sleep coach should keep in mind and utilize, found in this thesis are: functionality, interest, users' belief in own willpower, being in control, timing of reminders, content of reminders, perceived effort, perceived benefits, and ambivalence. For example, developers could design a virtual coach in such a way that adherence support is postponed until users need that help. Although this thesis focused on a virtual sleep coach for people suffering from insomnia, we expect that designers of other computerized health interventions might benefit from the gained insights as well.

6.4. Future research

As with all research the results of this thesis put forward new research questions. The limitations of the thesis could all be addressed by future research. So, future work could investigate virtual coaches on smartphone in other health domains, the effectiveness of various CBT-I components could be tested individually, cognitive therapy could be added to the app, the studies could be done in another culture, and other persuasive strategies could be tested. Additionally, the scope of the research could be shifted to integrating the app in the healthcare system (Andersson, 2010). With this shift new questions arise, and the following topics could then be investigated: cost-effectiveness of the app, the costs in terms of technical support to keep it running as well as the time it takes for therapists to support patients with the app, and the best format to integrate the app in the current healthcare system. For example, health apps can become stand-alone devices, or they can become part of a blended care approach. Additionally, the processes before and after CBT-I via an app should be studied. This thesis already made one small step towards this end-to-end process view, however, more work is needed to better understand and help people suffering from insomnia who are in the precontemplation, contemplation, or maintenance stage.

Other future work includes studying self-learning adaptive personalized virtual coaches. A personalized virtual coach could apply those persuasive strategies to which an individual user is susceptible. Kaptein (Kaptein, 2011) already started this research by creating personal persuasion profiles. Participants received a message based on

one of Cialdini's principles (Cialdini, 1993) with the aim to perform a certain behaviour. The behaviour was for instance taking the stairs instead of the elevator, or eating less candy bars. The effect of the message was measured by tracking the behaviour of the participants, and a personal persuasion profile was based on the measured effects. Simply stated, if a message was ineffective the score of that principle was lowered in the personal persuasion profile of that individual, and next time that person got a message based on another principle. If a message was effective the score of that principle was increased. Future work should expand on this idea and not only use messages based on Cialdini's principles, but include more persuasive strategies, like reminders, negotiation, rewards, etc. Personal profiling could also be applied to the therapy itself. Some people might benefit more from cognitive therapy as others, so this CBT-I component could then be offered earlier in the therapy. To have a starting point for personalization, people could fill in a general questionnaire before starting with the therapy. Included concepts could be, locus of control (Hertog, 1992), need for cognition (Cacioppo & Petty, 1982), and need for affect (Maio & Esses, 2001). Based on these personal characteristics a first educated guess could be made by the virtual coach on how to approach the user. Over time, the personal profile the coach holds on the user should be expanded and the coach his behaviour can be adjusted. Virtual coaches on a smartphone are very well suited to include personalisation because they are able to track certain behaviours relatively simple. In case of the virtual sleep coach, it is easy to track if a person filled in their diary in response to a reminder for example. Not only software sensors, also hardware sensors can be used to track behaviour. Currently, smartphones are equipped to measure level of activities (e.g., by GPS and accelerometers), which could be helpful for a coach that helps people to be more active. In the future, however, it is expected that smartphones will be able to measure much more, like temperature, humidity, heartbeat, and maybe even brain activity. All these sensors can be used to measure the effect of persuasive strategies and build up a personal persuasion profile.

In addition, virtual coaches as well as other intelligent systems should be able to cope with morality. Normative systems that formalise and implement policies should be studied urgently, since more and more decisions tasks will be automated. For example, the virtual sleep coach gives advice on bedtimes, often decreasing time in bed inducing sleep deprivation. Being sleep deprived can impair people to perform certain tasks, such as driving a car or operating heavy machinery. Therefore, the sleep coach never advised bedtimes less than 5 hours. Furthermore, warnings were given to the participants when their average sleep time dropped below the five hours, and the sleep restriction exercise was intermitted. The virtual sleep coach 'just' gave advice for now, but more advanced coaches could actually take decisions with possibly unforeseen consequences. Therefore intelligent systems should be 'aware' of their own limitations and communicate these to the users. In healthcare possible solutions might be in blended care or human monitoring.

6.5. Final remarks

Approximately 10% of the population is suffering from insomnia, which leads to significant impaired functioning during the day. The common treatment for insomnia is cognitive behavioural therapy (CBT-I), which traditionally is delivered by a therapist. However, self-help books are not uncommon either, and more recently self-help programs via the internet have become available, followed by self-help apps. Circa 6% of the total number of apps focus on mental health issues (Donker et al., 2013). Because health apps are relatively new, evaluations and standards have not been established yet, and are developed concurrently. So, the aim of this thesis was to study the effectiveness of a virtual sleep coach app and the way persuasive strategies could contribute to treatment adherence, since treatment adherence has been mentioned by numerous authors to be one of the main problems in CBT-I. The results show that an app can be effective in treating insomnia, and that adherence can, for example, be supported by reminders. However, adherence is still a topic of concern that should be studied more thoroughly.

6.6. References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211.

Ajzen, I., & Fishbein, M. (1988). Theory of reasoned action-Theory of planned behavior. *University of South Florida*.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed ed.). Arlington, VA: American Psychiatric Publishing.

Andersson, G. (2010). The promise and pitfalls of the internet for cognitive behavioral therapy. *BMC Medicine, 8*. doi: 10.1186/1741-7015-8-82

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review, 84*(2), 191.

Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Lichstein, K. L., & Morin, C. M. (2006). Recommendations for a standard research assessment of insomnia. *Sleep: Journal of Sleep and Sleep Disorders Research; Sleep: Journal of Sleep and Sleep Disorders Research*.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *J Pers Soc Psychol, 42*(1), 116.

Cialdini, R. B. (1993). Influence: Science and practice. *New York, NY: Harper Collins*.

Donker, T., Petrie, K., Proudfoot, J., Clarke, J., Birch, M.-R., & Christensen, H. (2013). Smartphones for smarter delivery of mental health programs: a systematic review. *J Med Internet Res, 15*(11).

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Fitrianie, S., Griffioen-Both, F., & Beun, R. (in preparation). Talk and Tools: The Best of Both Worlds in Mobile Interfaces.

Fogg, B. J. (2009). *A behavior model for persuasive design.* Paper presented at the Persuasive 2009, Claremont, California, USA.

Hertog, P. C. d. (1992). Instrumenteel onderzoek. De 'IE-18 locus of control'- vragenlijst: betrouwbaarheid en validiteit van een gewijzigde versie. *Nederlands tijdschrift voor de psychologie, 47*, 82-87.

Kaptein, M. (2011). *Personalized Persuasion in Ambient Intelligence.* Paper presented at the PhD Thesis.

Maio, G. R., & Esses, V. M. (2001). The need for affect: Individual differences in the motivation to approach or avoid emotions. *Journal of personality, 69*(4), 583-614.

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation Science, 6*(1), 42.

Proudfoot, J. G., Parker, G. B., Pavlovic, D. H., Manicavasagar, V., Adler, E., & Whitton, A. E. (2010). Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. *J Med Internet Res, 12*(5), e64.

Ripple, L. (1955). Motivation, capacity, and opportunity as related to the use of casework service: Theoretical base and plan of study. *Social Service Review, 29*(2), 172-193.

Rogers, E. M. (2010). *Diffusion of innovations*: Simon and Schuster.

Torous, J., Friedman, R., & Keshavan, M. (2014). Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. *JMIR mHealth and uHealth, 2*(1), e2.

Varnum, M. E., Grossmann, I., Kitayama, S., & Nisbett, R. E. (2010). The origin of cultural differences in cognition the social orientation hypothesis. *Current directions in psychological science, 19*(1), 9-13.

Appendices

A. Studies included in the meta-analysis

	Authors	Year	Title
1	Oosterhuis and Klip	1997	The treatment of insomnia through mass media, the results of a televised behavioural training programme
2	Rybarczyk et al.	2002	Efficacy of two behavioral treatment programs for comorbid geriatric insomnia
3	Ström, Petterson, Andersson	2004	Internet-based treatment for insomnia: a controlled evaluation
4	Suzuki et al.	2008	Evaluation of an internet-based self-help program for better quality of sleep among Japanese workers: a randomized controlled trial
5	Ritterband et al.	2009	Efficacy of an Internet-based behavioral intervention for adults with insomnia
6	Van Straten et al.	2009	Self-help treatment for insomnia through television and book: a randomized trial
7	Vincent and Lewycky	2009	Logging on for better sleep: RCT of the effectiveness of online treatment for insomnia
8	Riley, Mihm, Behar, Morin	2010	A computer device to deliver behavioural interventions for insomnia
9	Lancee et al.	2011	Internet-delivered or mailed self-help treatment for insomnia? A randomized waiting-list controlled trial
10	Ritterband et al.	2011	Initial evaluation of an internet intervention to improve the sleep of cancer survivors with insomnia
11	Espie et al.	2012	A randomized, placebo-controlled trial of online cognitive behavioural therapy for insomnia disorder delivered via an automated media-rich web application
12	Haimov and Shatil	2013	Cognitive Training Improves Sleep Quality and Cognitive Function among Older Adults with Insomnia
13	Lancee et al.	2013	Baseline depression levels do not affect efficacy of cognitive-behavioral self-help treatment for insomnia
14	Lancee et al.	2013	Motivational support provided via email improves the effectiveness of internet-delivered self-help treatment for insomnia: A randomized trial

	Authors	Year	Title
15	Lawson et al.	2013	Validating a mobile phone application for the everyday, unobtrusive, objective measurement of sleep
16	Van Straten et al.	2013	Guided Internet-delivered cognitive behavioural treatment for insomnia: a randomized trial
17	Holmqvist, Vincent, Walsh	2014	Web- vs telehealth-based delivery of cognitive behavioral therapy for insomnia: a randomized controlled trial
18	Lipschitz, Landward, Nakamura	2014	An exploratory study of an online mind-body program for poor sleepers in a community sample

Not included because of a lack of data:

Authors	Year	Title
Morawetz et al.	1989	Behavioral self-help treatment for insomnia: A controlled evaluation
Riedel et al.	1995	Sleep compression and sleep education for older insomniacs: Self-help versus therapist guidance
Chen et al.	2013	Enhancing adherence to cognitive behavioral therapy for insomnia through machine and social persuasion

B. General notes about the meta-analysis

The numbers are based on the people who participated in the study, not on the number of people that filled out a particular questionnaire. E.g. if adherence in Lancee (2011) is calculated only based on the people who actually filled out the adherence questionnaire, then the adherence rate would be 89/168=53%

If a study has more than one follow-up assessment, data of the first follow-up is used. This holds for the following studies
- Lancee (2011) 4 weeks, 8 weeks, and 48 weeks follow-up
- Lipschitz (2014) 1 week, 1 month, and 3 months follow-up
- Van Straten (2013) 4 weeks, and 18 weeks follow-up

The 19 people who adhered in Vincent (2009) is a hypothetical number, and is calculated as follows: Over several weeks adherence to different exercises was measured. The total times all the participants together could adhere to the exercises was 354 times. 116 times adherence was higher than a set threshold. These numbers are used to calculate the adherence percentage (116/354= 33%), which is then transformed into the hypothetical number 19 (0,33*59 = 19), which indicates how many people of the 59 participants have adhered hypothetical. Note that, the calculated numbers are based on the people who participated in the study. If adherence is calculated only based on the people who filled out the adherence questionnaire, then the adherence rate would be 116/193= 60%

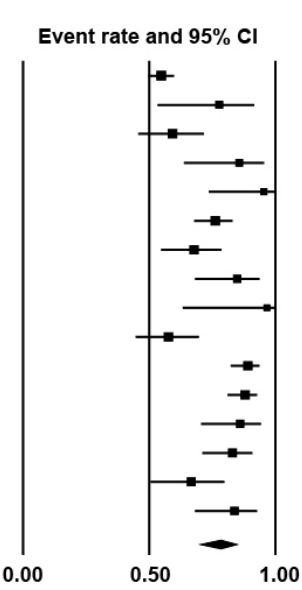
Lancee (2013, #13) consists of three groups of participants. Participants that suffer from high depression, mild depression and, low depression.

Lancee (2013, #14) compares CCBT-I with and without support.

C. Results of meta-analyses

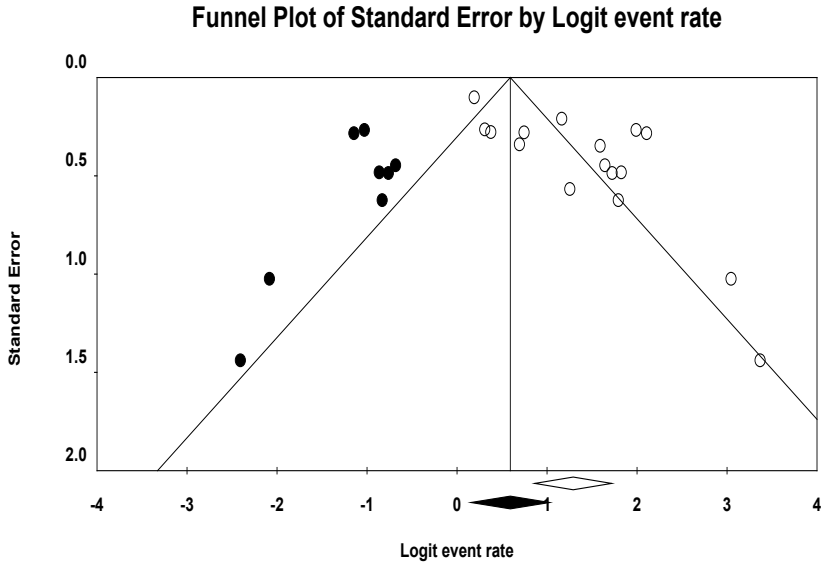
Experimental compliance post questionnaires

Study name	Raw data	Event rate	Lower limit	Upper limit	Relative weight
Oosterhuis (1997)	219 / 400	0.55	0.50	0.60	8.23
Rybarczyk (2002)	14 / 18	0.78	0.54	0.91	5.35
Strom (2004)	32 / 54	0.59	0.46	0.71	7.38
Suzuki (2008)	18 / 21	0.86	0.64	0.95	4.97
Ritterband (2009)	21 / 22	0.95	0.74	0.99	2.94
van Straten (2009)	96 / 126	0.76	0.68	0.83	7.78
Vincent (2009)	40 / 59	0.68	0.55	0.78	7.37
Riley (2010)	28 / 33	0.85	0.68	0.94	5.92
Ritterband (2011)	14 / 14	0.97	0.63	1.00	1.80
Haimov (2013)	34 / 59	0.58	0.45	0.70	7.46
Lancee (2013b) - With	115 / 129	0.89	0.83	0.93	7.34
Lancee (2013b) - Without	117 / 133	0.88	0.81	0.92	7.44
Lawson (2013)	31 / 36	0.86	0.71	0.94	5.94
van Straten (2013)	49 / 59	0.83	0.71	0.91	6.91
Holmqvist (2014)	26 / 39	0.67	0.51	0.80	6.96
Lipschitz (2014)	31 / 37	0.84	0.68	0.93	6.20
Random model		0.78	0.70	0.85	



Heterogeneity: $Q_{15} = 114.84$, $px < .001$, $I^2 = 86.94$, indicates substantial heterogeneity in the data, which supports the choice for a random-effects model.

Experimental compliance post questionnaires

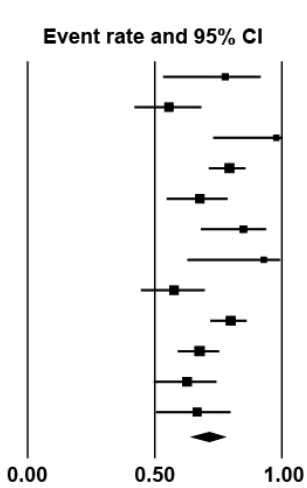


White dots indicate observed studies. The black dots indicate imputed data.

Publication bias: The funnel plot is noticeably asymmetric, with a majority of the smaller studies clustering to the right of the mean. This impression is confirmed by Egger's test (p = .002, two-tailed).

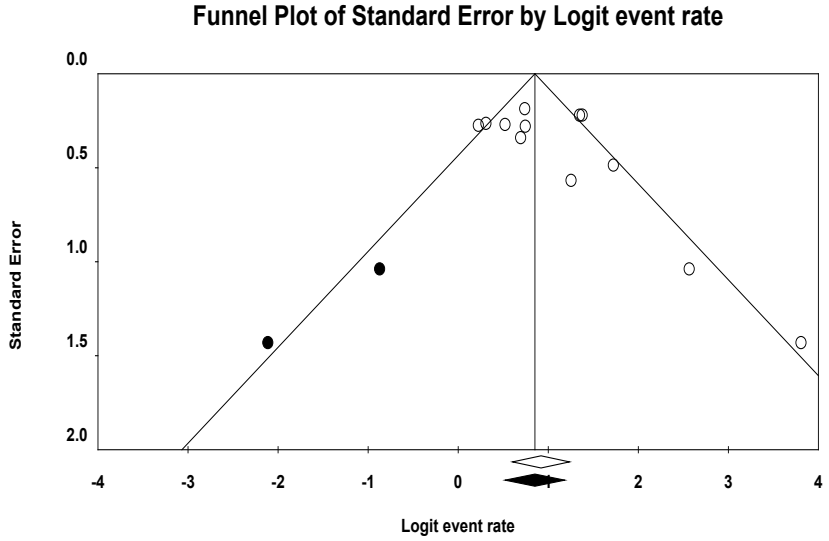
Experimental compliance post diaries

Study name	Raw data	Event rate	Lower limit	Upper limit	Relative weight
Rybarczyk (2002)	14 / 18	0.78	0.54	0.91	5.21
Strom (2004)	30 / 54	0.56	0.42	0.68	10.26
Ritterband (2009)	22 / 22	0.98	0.73	1.00	1.17
van Straten (2009)	100 / 126	0.79	0.71	0.86	11.46
Vincent (2009)	40 / 59	0.68	0.55	0.78	10.15
Riley (2010)	28 / 33	0.85	0.68	0.94	6.28
Ritterband (2011)	13 / 14	0.93	0.63	0.99	2.07
Haimov (2013)	34 / 59	0.58	0.45	0.70	10.49
Lancee (2013b) - With	103 / 129	0.80	0.72	0.86	11.47
Lancee (2013b) - Without	90 / 133	0.68	0.59	0.75	12.22
van Straten (2013)	37 / 59	0.63	0.50	0.74	10.36
Holmqvist (2014)	26 / 39	0.67	0.51	0.80	8.85
Random model		0.71	0.65	0.77	



Heterogeneity: $Q_{11} = 33.26$, $p < .001$, $I^2 = 66.92$, suggests that the data is heterogeneous and supports the choice for a random-effect model.

Experimental compliance post diaries

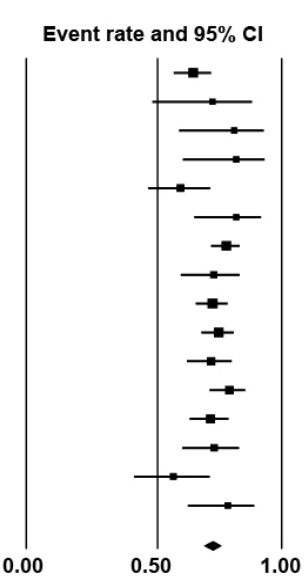


White dots indicate observed studies. The black dots indicate imputed data.

Publication bias: The shape of the funnel plot does not suggest significant publication bias, which is confirmed by Egger's test statistic, $p = 0.21$.

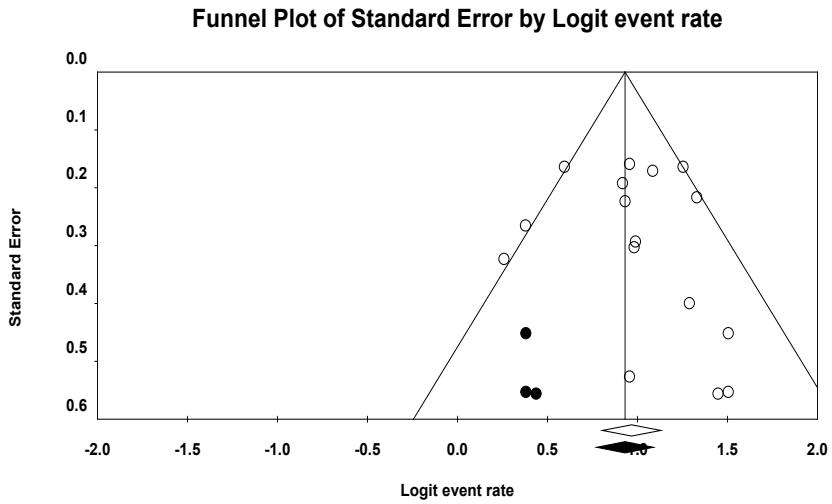
Experimental compliance follow-up questionnaires

Study name	Raw data	Event rate	Lower limit	Upper limit	Relative weight
Oosterhuis (1997)	105 / 163	0.64	0.57	0.71	10.35
Rybarczyk (2002)	13 / 18	0.72	0.48	0.88	2.17
Suzuki (2008)	17 / 21	0.81	0.59	0.93	1.97
Ritterband (2009)	18 / 22	0.82	0.60	0.93	1.99
Vincent (2009)	35 / 59	0.59	0.46	0.71	6.25
Riley (2010)	27 / 33	0.82	0.65	0.92	2.82
Lancee (2011)	168 / 216	0.78	0.72	0.83	10.35
Espie (2012)	40 / 55	0.73	0.60	0.83	5.23
Lancee (2013a) - Low	143 / 198	0.72	0.66	0.78	10.61
Lancee (2013a) - Mild	136 / 182	0.75	0.68	0.81	10.00
Lancee (2013a) - High	71 / 99	0.72	0.62	0.80	7.68
Lancee (2013b) - With	102 / 129	0.79	0.71	0.85	7.94
Lancee (2013b) - Without	95 / 133	0.71	0.63	0.78	8.98
van Straten (2014)	43 / 59	0.73	0.60	0.83	5.47
Holmqvist (2014)	22 / 39	0.56	0.41	0.71	4.77
Lipschitz (2014)	29 / 37	0.78	0.62	0.89	3.44
Random model		0.72	0.69	0.76	



Heterogeneity: $Q_{15} = 25.16$, $p = .048$, $I^2 = 40.39$, indicates heterogeneity, which support the choice for a random-effects model.

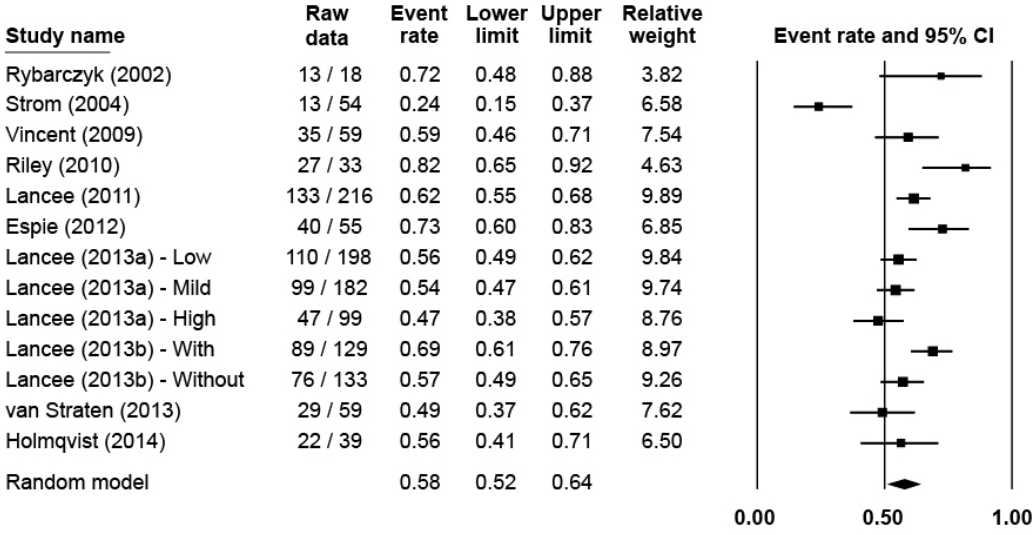
Experimental compliance follow-up questionnaires



White dots indicate observed studies. The black dots indicate imputed data.

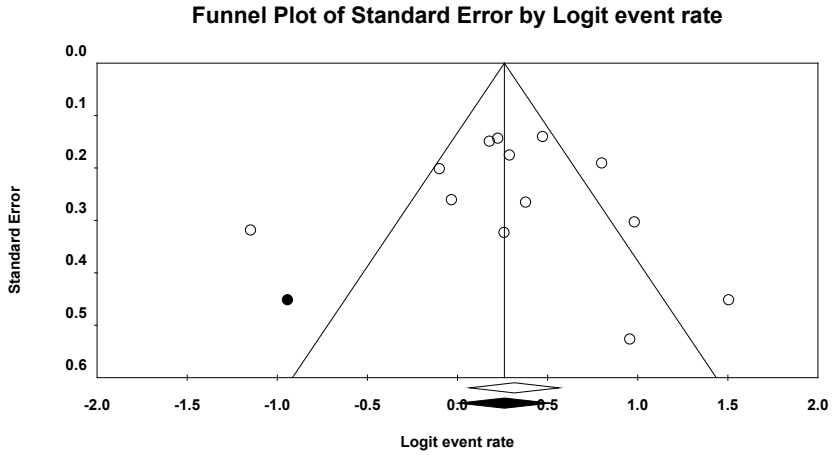
Publication bias: The shape of the funnel plot in did not reveal any indication of funnel plot asymmetry. This visual impression was also confirmed by Egger's test with $p = 0.61$, two-tailed.

Experimental compliance follow-up diaries



Heterogeneity: $Q_{12} = 49.54$, $p < .001$, $I^2 = 75.77$, indicates substantial heterogeneity and supports the choice for a random-effects model.

Experimental compliance follow-up diaries

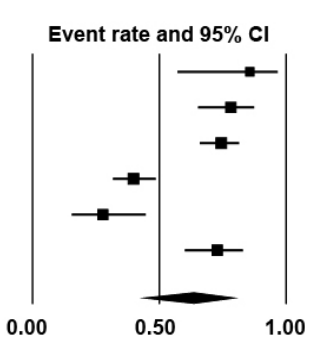


White dots indicate observed studies. The black dots indicate imputed data.

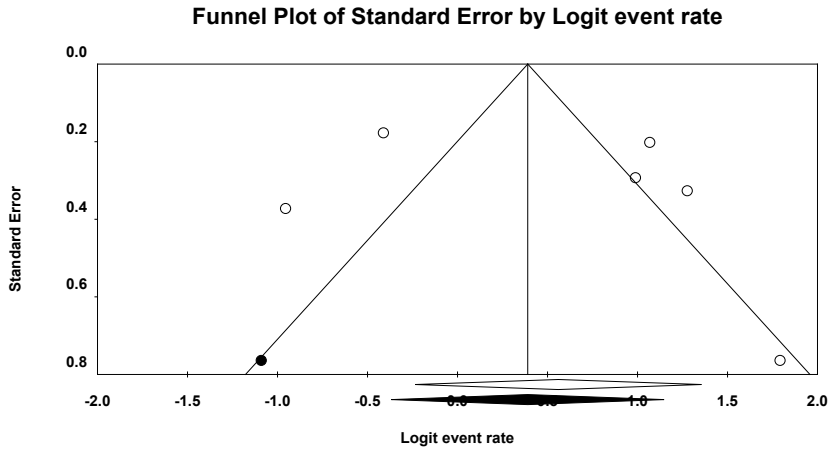
Publication bias: The shape of the funnel plot in did not reveal asymmetry. This visual impression was also confirmed by Egger's test with $p = 0.75$, two-tailed.

Logged treatment adherence

Study name	Raw data	Event rate	Lower limit	Upper limit	Relative weight
Ritterband (2011)	12 / 14	0.86	0.57	0.96	11.47
Espie (2012)	43 / 55	0.78	0.65	0.87	17.19
Lancee (2013b) - With	96 / 129	0.74	0.66	0.81	18.46
Lancee (2013b) - Without	53 / 133	0.40	0.32	0.48	18.66
Lawson (2013)	10 / 36	0.28	0.16	0.44	16.64
van Straten (2013)	43 / 59	0.73	0.60	0.83	17.57
Random model		0.64	0.44	0.79	



Heterogeneity: $Q_5 = 59.76$, $p < .001$, $I^2 = 91.63$, indicates that the data is heterogeneous, which supports the choice for a random-effects model.

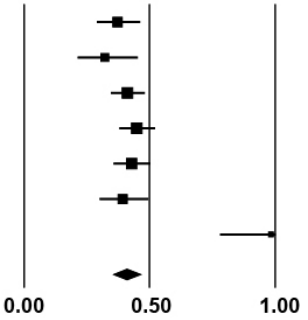


White dots indicate observed studies. The black dots indicate imputed data.

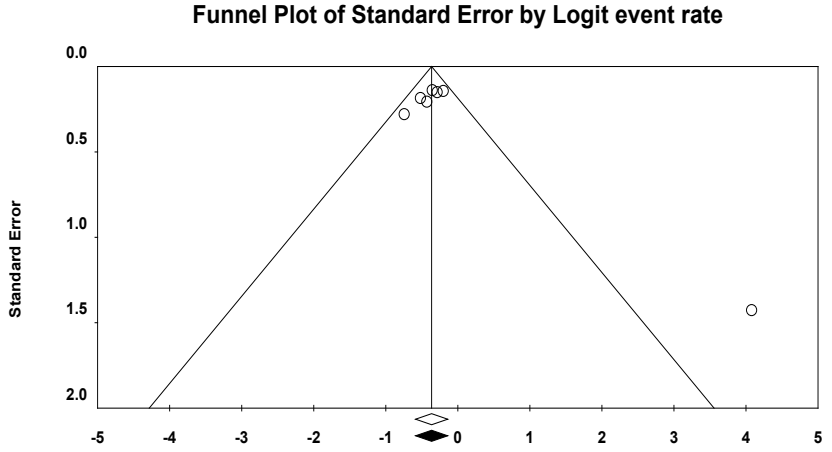
Publication bias: The shape of the funnel plot in did not reveal asymmetry. This visual impression was also confirmed by Egger's test with $p = 0.59$, two-tailed.

Self-reported treatment adherence

Study name	Raw data	Event rate	Lower limit	Upper limit	Relative weight
van Straten (2009)	47 / 126	0.37	0.29	0.46	16.16
Vincent (2009)	19 / 59	0.32	0.22	0.45	10.45
Lancee (2011)	89 / 216	0.41	0.35	0.48	19.84
Lancee (2013a) - Low	89 / 198	0.45	0.38	0.52	19.45
Lancee (2013a) - Mild	78 / 182	0.43	0.36	0.50	18.87
Lancee (2013a) - High	39 / 99	0.39	0.30	0.49	14.62
Lipschitz (2014)	29 / 29	0.98	0.78	1.00	0.62
Random model		0.41	0.36	0.46	



Heterogeneity: $Q_5 = 13.88$, $p = .031$, $I^2 = 56.77$, indicates heterogeneity and supports the choice for a random-effects model.



White dots indicate observed studies. The black dots indicate imputed data.

Publication bias: The shape of the funnel plot in did not reveal asymmetry. This visual impression was also confirmed by Egger's test with $p = 0.29$, two-tailed.

D. Precontemplation scenario

Leo downloaded the app yesterday. It appeared that the app suits his problems, so the program could help him, provided that he adhere to the exercise. During his breakfast Leo grabs his phone and he sees that he has a message from his new Sleep Coach App. He wonders what the app has to say. The coach wishes him good morning and indicates that there is a questionnaire for Leo before the sleep exercises start. Leo fills out the readiness to change questionnaire. It appears that Leo is in the precontemplation phase. The coach explains that motivation for change is an important success factor. Low motivation often leads to the non-completion of the therapy. From the questionnaire it seems that it might be inconvenient to immediately start with the sleep exercises. First, Leo and the coach will do some exercises to determine what Leo actually wants.

(precontemplation assignment 1)

The coach invites Leo to watch a movie of other people with sleeping problems. The coach shows a movie of peers who discuss their habits when they had sleeping problems. The peers were unaware of the fact that their old lifestyle have had a negative impact on their sleep quality.

(readiness to change measurement)

When the movie is finished the coach asks Leo whether he is willing to change his habits. Leo replies that he wants to sleep better, but he does not want to change his whole lifestyle. Because of his answer Leo gets another exercise belonging to the precontemplation stage. Leo cannot continue with sleep exercises, he is forced to do the motivational exercises first. Leo finishes his breakfast and he does not feel like performing another exercise. So, he closes the app, brushes his teeth and leaves for work.

Claims belonging to this scenario

1. At the start of the program, some users can be insufficiently motivated to complete the therapy.
2. Users who are not yet in the preparation phase, have a very high chance of dropping out.
3. Instead of convincing a user to change, the coach should support the thinking process about changing.
4. The user is only allowed to continue to the next readiness-to-change-phase, when the current phase is sufficiently completed.
5. Some users have not yet realized that their thoughts and habits are a problem for their sleep.

It is a good idea that the coach gives a warning that the user is not likely to complete the therapy due to a lack of motivation

E. Claims in Dutch and English

The claims discussed in different focus groups are listed below per scenario.

Scenario A: introduction with Leo

English	*Dutch*
<u>Potential users</u>	
I want to choose immediately between different coaches, this should not be done via settings.	Ik zou meteen willen kunnen kiezen tussen verschillende coaches, dit zou niet via de instellingen moeten gaan.
Different roles (such as the coach and professor) can best be performed by different virtual people.	Ik denk dat verschillende rollen (zoals de coach en de professor) het best uitgevoerd kunnen worden door verschillende virtuele personen.
<u>Coaches</u>	
If the user can choose between different coaches, that will ensure a better (trust) relationship between the user and the coach.	De keuze die de gebruiker kan maken tussen verschillende coaches, zorgt voor een betere (vertrouwens)band tussen de gebruiker en de coach.
Different roles (such as the coach and professor) can best be performed by different virtual people.	Verschillende rollen (zoals de coach en de professor) kunnen het best uitgevoerd worden door verschillende virtuele personen.

Scenario B: introduction with Marie

English	*Dutch*
<u>Sleep experts</u>	
An app which offers CBT-I is a good first step in a stepped care model.	Een app die CBT-I aanbiedt is een goede eerste stap in een stepped-care model.
One or more therapists, who can answer questions from users, should be connected to the app.	Er zouden één of meerdere therapeuten met de app verbonden moeten zijn, die vragen van gebruikers kunnen beantwoorden.
<u>Coaches</u>	
Splitting the therapy in small parts is a good idea.	Het opdelen van de therapie in kleine blokjes is een goed idee.
The interaction between coach and user (e.g. answering a question) improves the involvement of the user.	De interactie tussen de coach en de gebruiker (bijv. invullen/beantwoorden van een vraag) verhoogt de betrokkenheid van de gebruiker.
In the beginning, people are not inclined to give much information about themselves. The more users work with the system, the more willing they are to give information about themselves.	In het begin zijn mensen nog niet geneigd om veel informatie over zichzelf te geven. Hoe langer gebruikers met het systeem werken hoe meer bereid ze zijn om deze informatie over zichzelf te geven.
Personalization of the app based on user characteristics (established facts about the user, such as gender, name, family) ensures that the intervention will be used better than when there would be no personalization.	Personalisatie op basis van gebruikerskenmerken (vaststaande feiten over de gebruiker, zoals geslacht, naam, familie) van de app zorgt ervoor dat de interventie beter gebruikt wordt, dan wanneer er geen personalisatie zou plaats vinden.

Scenario C: introduction with Klaartje

English	*Dutch*
Potential users	
I think it's a good idea that I could indicate whether I would like more information or not.	Ik vind het een goed idee dat de ik zelf aan zou kunnen geven of ik nog meer informatie zou willen of niet.
To me it seems nice to know what the program does and what the coach does, shortly after the start of the program.	Het lijkt mij fijn om zo vroeg in het programma al kort te horen wat de coach doet en wat hij kan.
Getting in touch with other people who also follow the program would help me to adhere to the program longer.	In contact komen met andere mensen die het programma ook volgen, zou mij helpen om het langer vol te houden.
I think it's a good idea that the coach indicates how people around me could help me with certain assignments.	Ik vind het een goed idee dat de coach bij sommige opdrachten aangeeft hoe de mensen in mijn omgeving me zouden kunnen helpen bij opdrachten.

Scenario D: introduction new lifestyle with Marie

English	*Dutch*
Potential users	
Realistic expectations are raised by the explanation, therefore it is easier to complete the program.	Door de uitleg worden realistische verwachtingen geschept die er voor zorgen dat het gemakkelijker voor mij is het programma af te ronden.
Asking a question about the information given by the app, would make me remember the information better.	Het stellen van een vraag over de informatie die de app heeft gegeven, zou er voor zorgen dat ik de informatie beter onthoud.
Reading other users' stories could help me to keep using the app.	Het lezen van verhalen van andere gebruikers zou mij kunnen helpen om de app uiteindelijk langer te gebruiken.
If I do not know the purpose and the underlying principles of something, I will give up faster.	Als ik het doel en de achterliggend principes van iets niet kent, dan geef ik eerder op.
Giving a reward for answering a question right, would motivate me to read the information carefully.	Het geven van een beloning voor het goed beantwoorden van de vraag, zou er voor zorgen dat ik gemotiveerder ben om de informatie goed te lezen.
The choice to read or hear certain information again, doesn't need to be offered to me.	De keuze om de informatie nog een keer te kunnen lezen of horen, hoeft voor mij **niet** geboden te worden door het programma.
Coaches	
Asking a question about the information given by the app, would make users remember the information better.	Het stellen van een vraag over de informatie die de app heeft gegeven, zorgt ervoor dat gebruikers de informatie beter onthouden.
Giving a reward for answering a question right, would motivate people to read the information more careful and answer the question correctly.	Het geven van een beloning voor het goed beantwoorden van de vraag, zorgt ervoor dat mensen gemotiveerd zijn de vragen goed te beantwoorden en de informatie goed te lezen.

If people do not know the purpose and the underlying principles of something, they will give up faster.	Als je het doel en de achterliggend principes van iets niet kent, dan geef je eerder op.
Reading stories from peers can help users with their problems.	Het lezen van verhalen van peers kan gebruikers helpen met hun problemen.

Scenario E: inclusion/exclusion with Leo

English	*Dutch*
Sleep experts	
Excluding people from the program can be done later. This moment is too early.	Het onderscheid maken tussen de doelgroep en mensen die daarbuiten vallen kan later in het programma gebeuren. Dit moment is te vroeg.
The choice to read or hear certain information again, doesn't need to be offered by the program.	De keuze om de informatie nog een keer te kunnen lezen of horen, hoeft niet geboden te worden door het programma.
It is a good idea to have the possibility to directly schedule an appointment with a therapist, if a questionnaire indicates that there could be a problem other than insomnia.	Een mogelijkheid om direct een afspraak met een therapeut te maken als uit de vragenlijst blijkt dat er een ander probleem dan insomnie zou kunnen zijn is een goed idee.
For a therapist it is useful that the results of a questionnaire are saved, so that the therapist can assess the data.	Voor een therapeut is het handig als de uitslag van een vragenlijst wordt opgeslagen, zodat de therapeut dat kan bekijken.

Scenario F: Precontemplation phase with Leo

English	*Dutch*
<u>Sleep experts</u>	
When starting the therapy some people might be insufficiently motivated to finish it.	Sommige personen kunnen bij het begin van het gebruik van de app, onvoldoende gemotiveerd zijn om de therapie te doorlopen.
The user may only continue with the next "readiness-to-change-phase", when the current phase is completed. The coach decides on this.	De gebruiker mag pas verder naar de volgende "bereidheid-tot-verandering-fase", als de huidige fase voldoende is afgerond. De coach beslist dit.
It is good that the coach gives a warning if it is likely that the user will not complete the therapy due to a lack of motivation.	Het is goed dat de coach een waarschuwing geeft dat de kans groot is dat de gebruiker de therapie waarschijnlijk niet zal afronden door een gebrek aan motivatie.
Instead of convincing a user, the coach should support the thinking processes of the user.*	In plaats van overtuigen moet de coach het denkproces omtrent de bereidheid tot verandering van een gebruiker ondersteunen.*
<u>Coaches</u>	
It is good that the coach gives a warning if it is likely that the user will not complete the therapy due to a lack of motivation.	Het is goed dat de coach een waarschuwing geeft dat de kans groot is dat de gebruiker de therapie waarschijnlijk niet zal afronden door een gebrek aan motivatie.
Users who are still in the preparation phase, have a high chance to stop early with the program.	Gebruikers die nog voor de preparation fase zitten, hebben een hele hoge kans om vroegtijdig met het programma te stoppen.
It is a good idea that if the user is not sufficiently motivated, the user will be offered a different but equal exercise, in order to increase the level of motivation.	Het is een goed idee dat als de gebruiker nog onvoldoende gemotiveerd is er een gelijkwaardige, maar andere oefening wordt aangeboden, om de gebruiker alsnog op het gewenste motivatieniveau te krijgen.

* These claims were not discussed due to time limitations.

Scenario G: Contemplation phase with Klaartje

English	*Dutch*
Potential users	
I think it is unnecessary that the coach indicates how long an exercise will approximately take in advance.	Ik vind het overbodig dat de coach van te voren aangeeft hoe lang een opdracht ongeveer gaat duren.
Giving compliments would motivate me to continue to do the exercises.	Het geven van complimenten, zou mij motiveren om door te gaan de opdrachten uit te voeren.
I would like to have the possibility to not know something, this would give a feeling of safety.	Ik zou het fijn vinden dat de mogelijkheid om iets niet weten gegeven wordt door de slaapcoach, hierdoor zou ik me veilig en serieus genomen voelen.
If I can choose the reminder time myself, I would respond better to that notification.	Als ik de tijd van een herinnering zelf kan kiezen, dan zou ik beter op die herinnering reageren.
I think it's a good idea that the coach would help me with planning an assignment, if I am not able to do the exercise right now.	Ik vind het een goed idee dat de coach mij zou helpen met het plannen van een opdracht, als ik deze niet op dit moment zou kunnen uitvoeren.
I would like it if the app sends me reminders for exercises etc.	Ik zou het fijn vinden als de app herinneringen stuurt voor het uitvoeren van opdrachten etc.
A user forum would not provide me support for doing the exercises.	Een forum met andere gebruikers zou mij **geen** ondersteuning bieden bij de opdrachten.
Coaches	
A user forum would support people with doing the exercises.	Een forum met andere gebruikers biedt goede ondersteuning bij de opdrachten.
Giving compliments motivates users to continue with the exercises.	Het geven van complimenten motiveert de gebruiker om door te gaan met de opdrachten.

By showing how far users are with an assignment (progress indicator), they are more motivated to continue.

Door te laten zien hoe ver gebruikers al zijn met een opdracht (voorgangsindicatie), zijn ze meer gemotiveerd om door te gaan.

It would give users a feeling of being taken seriously and safety, when the possibility to not know something is given.

Door gebruikers de mogelijkheid te geven dat ze iets niet weten, voelen ze zich veilig en serieus genomen

Scenario H: preparation phase with Marie

English	*Dutch*
<u>Potential users</u>	
I would like it if the app sends me reminders for exercises etc.	Ik zou het fijn vinden als de app herinneringen stuurt voor het uitvoeren van opdrachten etc.
I would not mind if the coach measures my motivation level and decides (on behalf of me) if I am motivated enough to go on to the next “readiness-to-change-phase”.	Ik zou het **<u>niet</u>** erg vinden dat de coach mijn motivatieniveau meet en besluit (in plaats van ik zelf) of ik genoeg gemotiveerd ben om door te gaan naar de volgende “bereidheid-tot-verandering-fase”.
Giving examples would be superfluous for me. I would not use the examples to do the exercises in a better way.	Het geven van voorbeelden zou voor mij overbodig zijn. Ik zou de voorbeelden niet gebruiken om de opdrachten beter uit te voeren.
Making a list of people around me who could help me with the intended changes would help me with the process of change.	Een lijstje maken van mensen in mijn omgeving die me zouden kunnen helpen met de beoogde verandering zou verandering voor mij vergemakkelijken.
<u>Coaches</u>	
If a user has not used the app for a long time, the app will send a reminder.	Als een gebruiker de app lang niet gebruikt, dan stuurt de app een herinnering.
By giving examples, users have a better idea of what to do and therefore can do the exercises in a better way.	Door het geven van voorbeelden, hebben gebruikers een beter idee van wat ze moeten doen en kunnen daardoor de opdrachten beter uitvoeren.
It is a good idea to make people think about how other people can help them with their intended change.	Het is goed om mensen na te laten denken over hoe anderen mensen hun kunnen helpen met de verandering die ze willen ondergaan.
It is better that the coach measures the motivation of the user and decides whether the user is motivated enough to continue to the next “readiness-to-change-phase”, than when the users decide this for themselves.	Het is beter dat de coach het motivatieniveau meet en besluit of de gebruiker genoeg gemotiveerd is om door te gaan naar de volgende “bereidheid-tot-verandering-fase”, dan dat de gebruiker dit zelf mag beslissen.

Scenario I: filling out the sleep diary with Leo

English	*Dutch*
<u>Potential users</u>	
Immediately starting with filling in the sleep diary (at the beginning of the use of the app), would increase my motivation to keep using the program.	Meteen starten (bij het begin van het gebruik van de app) met het invullen van het dagboek zou mijn motivatie om door te gaan met het programma verhogen.
To me it seems nice to get a reminder later on the day, when I did not complete the sleep diary yet.	Het lijkt mij prettig om later op de dag een herinnering te krijgen als ik het dagboek nog niet heb ingevuld.
Getting points would result in me doing the exercises better.	Het krijgen van punten zou er voor zorgen dat ik de opdrachten beter uitvoer.
Personalization of the app based on my preferences (such as type of coach, preferred reminder type) would make me use the app better, then if there was no personalization.	Personalisatie van de app op basis van mijn voorkeuren (zoals type coach, voorkeur voor remindertype) zou er voor zorgen dat ik de app beter zou gebruiken, dan wanneer er geen personalisatie zou plaats vinden.
<u>Sleep experts</u>	
Immediately starting with filling in the sleep diary (at the beginning of the use of the app), would increase users' motivation to keep using the program.	Meteen starten (bij het begin van het gebruik van de app) met het invullen van het dagboek verhoogd de motivatie van gebruikers om door te gaan met het programma.
It is a good idea to send a reminder later on the day, when a user did not complete the sleep diary yet.	Het is een goed idee dat de gebruiker later op de dag een herinnering krijgt als diegene het dagboek nog niet heeft ingevuld.
It is better to give no feedback to the users about their sleep during the first week. This means the user can not see any graphs about their sleep.	Het is beter om gebruikers de eerste week geen terugkoppeling te geven over hun slaap. Dit betekent dat de gebruiker geen grafiekjes en overzichtjes kunnen zien over hun slaap.
It is diagnostically important to fill in the diary immediately after someone wakes up.	Het is diagnostisch belangrijk om het dagboek in te vullen meteen nadat men wakker wordt.

Coaches

Getting points makes people do the exercises better.	Het krijgen van punten zorgt ervoor dat mensen de opdrachten beter uitvoeren.
On the basis of how well users fill in the diary, the willingness to change can be established.	Aan de hand van hoe goed gebruikers het dagboek invullen, kan de bereidheid tot verandering gemeten worden.
Personalization of the app based on user preferences (such as type of coach, preferred reminder type) would improve the app usage, compared to no personalization.	Personalisatie op basis van gebruikersvoorkeuren (zoals type coach, voorkeur voor herinneringstype) van de app zorgt ervoor dat de interventie beter gebruikt wordt, dan wanneer er geen personalisatie zou plaats vinden.

Scenario J: Setting goals with Marie

English	*Dutch*
Sleep experts	
It is a good idea that users can specify whether they want more information or not.	Het is een goed idee dat de gebruikers zelf aan kunnen geven of ze nog meer informatie willen of niet.
Not filling in the diary for one day, is not a problem.	Één dag het dagboek niet invullen is geen probleem.
Letting the program calculate the sleep efficiency is a good idea.	Slaapefficiëntie laten uitrekenen door het programma is een goed idee
It is better to first fill in the diary, and set goals for yourself afterwards than vice versa.	Het is beter om eerst het dagboek in te vullen en daarna doelen voor je zelf te stellen i.p.v andersom.
Coaches	
It is better to first fill in the diary, and set goals for yourself afterwards than vice versa.	Het is beter om eerst het dagboek in te vullen en daarna doelen voor je zelf te stellen i.p.v andersom.
It is ok when some goals are determined by the sleep coach.	Het is goed als sommige doelen door de slaapcoach worden bepaald.
A goal should be realistic and measurable (i.e. something with numbers).	Een einddoel moet realistisch en meetbaar zijn (dus iets met getallen).

Scenario K: The plan of change with Klaartje

English	*Dutch*
<u>Sleep experts</u>	
The user does not have the knowledge to develop an effective change plan, therefore it is better that the coach gives a suggestion.	De gebruiker heeft niet de kennis om een effectief veranderplan op te stellen, daarom is het beter als de coach een suggestie geeft.
It must be possible for the user to adapt the treatment plan.	Het moet voor de gebruiker wel mogelijk zijn om zelf het behandelplan aan te passen.
For the user to be in control, and therefore more motivated, is less important than a good therapy. In other words, a user can better adhere to half of a good therapy, then fully complete a bad therapy.	Dat de gebruiker in controle is en daardoor ook gemotiveerder is, is minder belangrijk dan een goede therapie. Anders gezegd, een gebruiker kan beter een halve goede therapie doorlopen, dan een slechte therapie volledig afronden.
Sharing the therapy progress is not ok.	Het delen van de voortgang van de therapie is niet oké.
<u>Coaches</u>	
It must be possible for the user to adapt the treatment plan.	Het moet voor de gebruiker mogelijk zijn om zelf het behandelplan aan te passen.
It is a good idea to have the user give explicit consent (for example, by putting a signature on the plan).	De gebruiker expliciete instemming laten geven (bijvoorbeeld door een handtekening te laten zetten onder het plan), is een goed idee.
By sharing the plan with others, such as friends, family and other people with sleep problems, users are more likely to execute the plan.	Door het plan te delen met anderen, zoals vrienden, familie en andere mensen met slaapproblemen, zijn gebruikers eerder geneigd het plan uit te voeren.
Sharing the therapy progress is not ok.	Het delen van de voortgang van de therapie is niet oké.

Scenario L: Therapy with Marie

English	*Dutch*
<u>Potential users</u>	
If I get a notification from the coach, I would like to have the possibility to snooze the reminder (repetition of the notification over 9 minutes), or to delete the notification completely.	Ik vind het een goed idee dat als er een melding van de coach komt, ik zou kunnen kiezen om de melding te sluimeren (herhaling over 9 min), of om de melding helemaal weg te drukken.
I think it's a good idea that when I turn the coach off, the app will still send a notification after 2 or3 hours.	Ik vind het een goed idee dat als ik de coach definitief wegdrukt, de app na 2 à 3 uur toch terug komt met de melding.
The explanation of an exercise should always be shown before the exercise starts.	Ik vind dat de uitleg van de opdracht altijd te zien zou moeten zijn voordat de opdracht begint.
I want to be able to click quickly through the information if I am already familiar with the information.	Ik zou snel door de informatie heen willen kunnen klikken als ik al bekend ben met de informatie.
<u>Sleep experts</u>	
The successive steps, education, instruction, execution, evaluation and prevention of relapse, are the best way to offer a therapy exercise.	De opeenvolgende stappen onderwijs, instructie, uitvoering, evaluatie en voorkomen van terugval, zijn de beste manier om een therapie opdracht aan te bieden.
The users should be able to click quickly through the information if they are already familiar with the information.	De gebruikers zouden snel door de informatie heen moeten kunnen klikken als ze al bekend zijn met de informatie.
Sometimes the section 'education' can be skipped.	Soms kan het onderdeel 'onderwijs' overgeslagen worden.
The instruction belonging to an exercise should always be shown before the exercise starts.	De ´instructie van´ de opdracht zou altijd te zien moeten zijn voordat de opdracht begint.
<u>Coaches</u>	
Monitoring motivation weekly, is not a good idea.*	Motivatie per week monitoren is **<u>geen</u>** goed idee.*

* These claims were not discussed due to time limitations.

The users should be able to click quickly through the information if they are already familiar with the information.*	De gebruikers zouden snel door de informatie heen moeten kunnen klikken als ze al bekend zijn met de informatie.*

* These claims were not discussed due to time limitations.

F. Short summary of personas

They were:

Marie, a 33 year old married mother of two young children, who is in preparation phase and got the app from her general practitioner

Klaartje, a 62 year old widow, who is in contemplation phase and got the app from a friend

Leo, a 49 year old married man, with three children studying, who is in precontemplation phase.

G. Measurement overview

Pre-measures	**Week 1,2, and 3**	**Post-measures**
Insomnia Severity	Objective adherence	Insomnia Severity Index
GSM usage ability	Subjective adherence	Ability to perform activity
Behavioural intention	Satisfaction with adherence	Score for the reminders
Locus of Control	Easy to initiate	Ranking of the reminders
	if reminder	*UTAUT:*
	Score/grade for reminder	Utility
	Opportunity	Effort
	Control	Social influence
	Predictability	Facilitating conditions
	Commitment	Attitude
	Motivation Diary	Self-efficacy
	Motivation Relax	Anxiety
	Irritation	Trust
	Remarks*	Behavioural intention

All measures were used in the expectation maximisation algorithm to fill in missing data, except the remarks denoted by *.

H. Pre-treatment questionnaire

GSM ability

I know how to use my smartphone
I know how I can respond to notifications on my smartphone
Other apps that I use, send me reminders sometimes
I regularly set reminders myself using my smartphone

Behavioural intention

The theory of planned behaviour states that behavioural intention predicts behaviour [40]. Therefore, behavioural intention (BI) was measured using the six questions below.
I plan to use the app for 3 weeks
I will follow the instructions/advice from the app
I plan to complete my sleep diary every day
I will definitely look at the overview of my sleep data
I am planning to do the relaxation exercise twice a day
If I have a question about the app, I will search for an answer

Locus of control

Locus of control (LoC) was measured via a 18-item Dutch questionnaire[1]. A higher internal locus of control has been found to influence diary adherence in an online lifestyle diary[2]. Higher scores indicate a higher external locus of control.

[1] Hertog, P. C. d. (1992). Instrumenteel onderzoek. De 'IE-18 locus of control'- vragenlijst: betrouwbaarheid en validiteit van een gewijzigde versie. Nederlands tijdschrift voor de psychologie, 47, 82-87.

[2] Blanson Henkemans OA, van der Boog PJM, Lindenberg J, van der Mast CAPG, Neerincx MA, Zwetsloot-Schonk BJHM. An online lifestyle diary with a persuasive computer assistant providing feedback on self-management. Technology and Health Care. 2009;17(3):253-67.

I. Weekly questionnaire

Adherence

Diary

How many times did you fill in the diary last week? If you don't know it exactly, estimate it to your best ability
Why did you not fill in the diary on some days?
I am satisfied with how often I have completed the diary last week
How many reminders did you get the past week about filling in the diary? If you are not sure, try to estimate it.

Relaxation

How many times did you do the relaxation exercise last week? If you don't know it exactly, estimate it to your best ability
Why did you not do the relaxation exercise on some days?
I am satisfied with how often I have done the relaxation exercise last week.
How many reminders did you get the past week about doing the relaxation exercise? If you are not sure, try to estimate it.

General

On a scale from 1 to 10, with 1 being the lowest and 10 being the highest rating. What grade would you give this kind of reminder?

Easy to initiate

Easiness to use was measured with four 7-point Likert scale statements. If an activity is hard to integrate in people's daily life, the probability that people will perform the activity decreases, since people's behaviour are affected by the principle of least effort [43].

Diary

It was hard to make time to fill in the diary
Filling in the diary was kind of a habit for me

Relaxation

The relaxation exercises were easy to integrate into my daily routines
I had to put in a lot of effort to not forget to do the relaxation exercise

Motivation

For measuring motivation the Situational Motivation Scale (SIMS)[1] was used.

[1] Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). Motivation and emotion, 24(3), 175-213.

J. Weekly Reminder questionnaire

Opportunity

Diary

The reminders for the diary arrived at inopportune times.
The reminders for the diary were timed well.
I always responded to the reminders of the diary.
Cronbach's alpha's: Self-set: .443, COM-B: .718

Relaxation

The reminders for the relaxation exercise were sent at the right time.
I often dismissed the reminder for the relaxation exercise, because it the time was inappropriate.
Reminders for the relaxation exercises came at times that did not suit me.
Cronbach's alpha's: Self-set: .776, COM-B: .789

Self-Empowerment

Control - Diary

I think that I had enough influence on the reminders for the diary
I had enough control over the reminders for the diary
I had no control over the reminders for the diary
Cronbach's alpha's: Self-set: .944, COM-B: .861

Control - Relaxation

I did not have enough control over the reminders for the relaxation exercise
I was in control regarding the reminders for the relaxation exercise
I could not influence the reminders for the relaxation exercise enough
Cronbach's alpha's: Self-set: .867, COM-B: .923

Predictability - Diary

I was not able to predict the reminder times for the diary
Reminders for the diary came unexpectedly
The reminders for the diary came at predictable times
Cronbach's alpha's: Self-set: .636, COM-B: .711

Predictability - Relaxation

I regularly wondered when the reminders for the relaxation exercises would come
The reminders for the relaxation exercises came at moment that I expected them to come
I think the reminders for the relaxation exercises arrived at predictable moments
Cronbach's alpha's: Self-set: .987, COM-B: .824

Commitment - Diary

I felt uncomfortable ignoring the reminders for the diary
When I acted on the reminders for the diary I felt content
I felt guilty when I did not respond to the reminders for the diary
Cronbach's alpha's: Self-set: .527, COM-B: .797

Commitment - Relaxation

I had the feeling I did not stick. to an agreement if I ignored the reminders for the relaxation exercises
I owned it to myself to follow the reminders of the relaxation exercise
I did not have any trouble ignoring the reminders for the relaxation exercises
Cronbach's alpha's: Self-set: .629, COM-B: .616

Irritation

Irritation was measured with six 7-point Likert scale statements. If people were irritated by the reminders the chance they will adhere decreases[1].

Diary

I got to many reminders for the diary
I appreciated the reminders for the diary
I was annoyed by the reminders for the diary
Cronbach's alpha's: Self-set: .627, COM-B: .744

Relaxation

I think the reminders for the relaxation exercises are nice
I got mad with the reminders of the relaxation exercises
I got to many reminders for the relaxation exercises
Cronbach's alpha's: Self-set: .688, COM-B: .809

[1] Bickmore T, Mauer D, Crespo F, Brown T. Persuasion, task interruption and health regimen adherence. Persuasive Technology: Springer; 2007. p. 1-11

K. Final questionnaire – users' experiences

Ability

I found it easy to fill in the diary
I thought it was difficult to perform the relaxation exercise
I totally understood the instructions of the relaxation exercise
I did not understand the instructions of the sleep diary

Satisfaction

Reminder preference: You have received three types of reminders; no reminders; self-set reminders; automatic reminders. Indicate which reminder you preferred. Start on the top with your favourite reminder and end with your least favourite reminder.

Appreciation reminders: On a scale from 1 to 10, in which 10 is the highest score, which scores would you give the reminder types below?

- No reminder
- Self-set reminder
- Automatic reminder

Appreciation app components: On a scale from 1 to 10, in which 10 is the highest score, which scores would you give the components below?

- The sleep diary
- The relaxation exercise
- The app in total

UTAUT

The users' experiences measure was based on the Unified Theory of Acceptance and Use of Technology (UTAUT)[1]. The Unified Theory of Acceptance and Use of Technology (UTAUT) defines eight concepts that measure technology acceptance and link them to intention and usage, and thereby possibly explain adherence. In addition, trust has been added to the UTAUT model as a predictor, since a lack of trust could negatively influence usage[2,3].

[1] Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: Toward a unified view. MIS quarterly. 2003:425-78
[2] Cheung C, Lee MK. Trust in Internet shopping: A proposed model and measurement instrument. AMCIS 2000 Proceedings. 2000:406
[3] Gefen D, Karahanna E, Straub DW. Trust and TAM in online shopping: an integrated model. MIS quarterly. 2003;27(1):51-90

Utility/Effect

With the app I could track my sleep pattern very well
By using the app I could detect problems in my sleep pattern
Using the app improved my daily quality of life
Cronbach's alpha: .712

Effort

It was easy for me to figure out how the app worked
The app was easy to use in daily life
The app was complicated, therefore it was hard for me to understand
Cronbach's alpha: .724

Social Influence

People who are important to me think I should use the app
My family supported me in using the app
My friends think I should use the app
Cronbach's alpha: .799

Facilitating Conditions

The app was not compatible with other products I use
I had the knowledge necessary to use the app
Someone was available for assistance with app difficulties
Cronbach's alpha: .549

Attitude

Using the app was a good idea
I liked using the app
I hated using the app
Cronbach's alpha: .751

Self-efficacy

I could use the app without any help
Using the app went well as long nothing unexpected happened
If there was no one around to help me, I preferred not to use the app
Cronbach's alpha: .871

Anxiety

The app was somewhat intimidating to me
It scared me to think that I could lose information by pressing the wrong button
I think the app could invade my privacy
Cronbach's alpha: .588

Trust

I trusted the information the app gave me
I felt distressed using the app
I have confidence in the app working well
Cronbach's alpha: .600

Behavioural Intention

I was determined to insert information into to app at the right time
I intended to look at the graphs of my sleep
My intention was to use the app for 3 weeks
Cronbach's alpha: .678

L. Interview questions

Why did you use the app?

Did you have any specific goals?

What do you think of the app?

How often did you plan on filling in the diary? And were you successful with that? / How well / often did you fill in the diary? / Every day or did you sometimes not fill it in?

If you filled out the diary when and where you did you usually do it?

Were there certain reasons (obstacles / things / events), why it sometimes did not work out to fill in the diary? If so, what were those obstacles?

How well did filling in your diary fit in your daily life?

Did the different types of reminders affect whether or not you filled out the diary?

How often did you plan to do the relaxation exercise?

And were you successful with that? / How well / often did you do the relaxation exercise? / Every day or did you sometimes not do it?

If you did the relaxation exercise when and where you did you usually do it?

Were there certain reasons (obstacles / things / events), why it sometimes did not work out to do the relaxation exercise? If so, what were those obstacles?

How well did the relaxation exercise fit in your daily life?

Did the different types of reminders affect whether or not you did the relaxation exercise?

Did you look at your sleep data? If so, when did you look at it? What exactly did you want to know / see? Did you find it? What do you think about the way the data is displayed?

To what extent was the app beneficial and effective?

What effect had the diary in your sleep?

What effect had the relaxation exercise on your sleep?

What kind of effect did the app have on yourself?

Did the app had an effect on something else in your life?

Were there any irritations regarding the app?

Did it cause irritation when you received reminder, or did you thought it was fine? What did you think about the reminders? What kind of reminder irritated you the most?

What would you like to see improved in the app? And what else?

What properties would you like to add to the app?

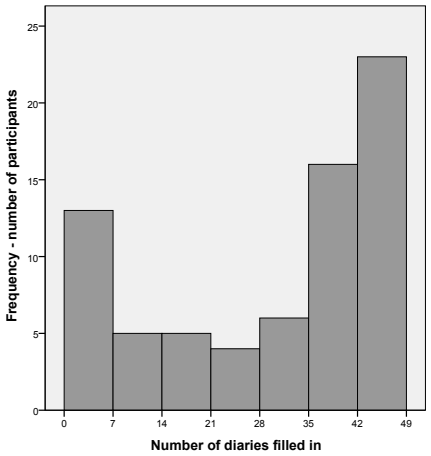
What can be added to the app, so you would adhere (even) longer or better?

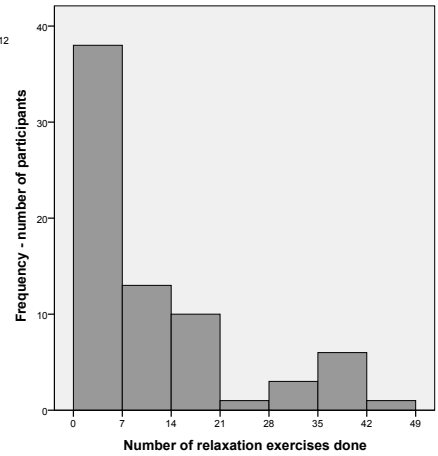
Would you recommend the App to others? Why?/ Why not?

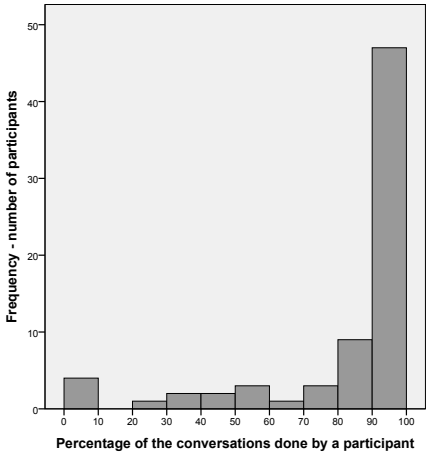
What did you think of the experiment itself, not the app, but everything else, such as emails, surveys, downloading, etc.?

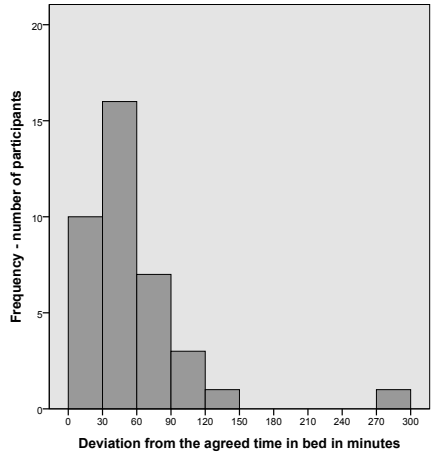
Are there other things you want to say about the app that I have not covered?

M. Adherence patterns









Dankwoord

Finally I am writing the last part of my thesis! It would not have been possible for me to write this thesis all by myself. I am really grateful for the help I got from the committee members, my supervisors, colleagues, friends, and family. Without you, this thesis would not have been possible.

I would like to thank all the committee members: Catholijn Jonker, Huib de Ridder, Harri Oinas-Kukkonen, Joyce Westerink, en Robbert Jan Beun. I appreciate the time you all took to read my dissertation and place comments where necessary. I am looking forward to the questions and discussions during the defence.

Willem-Paul, ik heb je bijna wekelijks gesproken en ik heb van onze samenwerking genoten. Onze praktische aanpak, open- en eerlijkheid heb ik altijd gewaardeerd. Soms verschilden we van visie, maar we kwamen er altijd samen uit. Bedankt voor alles. Mark, bedankt voor je wijsheid en rust die je me gaf als ik weer eens in sneltreinvaart dingen wilde uitvoeren. Met zijn drieën vormden we een goed team.

Ik heb het geluk gehad om binnen het Sleepcare team samen te mogen werken met prachtige mensen. Robbert Jan bedankt voor je enthousiasme over al onze ideeën en dat je ons aan 'je kindje' hebt laten werken. Jaap, ik heb veel van je geleerd, je was een soort onofficiële begeleider voor me. Bedankt voor je snelle antwoorden, je hoge kwaliteit feedback, en geruststelling die je me gaf in de soms chaotische praktijk. Fiemke, bedankt voor de vrolijke noot die je bracht en je inzet voor het project. Ik weet niet of we het zonder je hadden gered. Sandor, ook zonder jou hadden we de eindstreep niet gehaald. Bedankt voor het programmeren van de app. Ook al begrepen we elkaar niet altijd, we hadden wel altijd een gezamenlijk doel: de Sleepcare app. Rogier, Siska, en Reinder, ik heb voornamelijk met jullie vergaderd, maar dit was alles behalve saai. Ook jullie hebben het project een warm hart toegedragen en daarmee ook mij en dit proefschrift. Bedankt. Anna, helaas was het Sleepcare project toch niet helemaal wat je zocht, maar ik ben blij dat ik je heb leren kennen en ik wil je bedanken voor het warme onthaal in het begin van het proces. Het was fijn om ons samen af te vragen waar we in beland waren.

De meeste tijd heb ik doorgebracht in Delft, waar ik ook veel steun heb gehad van mijn collega's. Bedankt, Alex, Chris, Christian, Christina, Ding, Dwi, Dylan, Iris, Maaike, Marieke, Reyhan, Timea, Thomas, Vanessa, Vincent, Wenxin, Myrthe, Iris, en Ursula. Om in Delft te komen moest ik heel wat ritjes maken en die werden eerst een stuk gezelliger door Iris en daarna door Ursula. Iris bedankt voor je luisterende oor, je verhalen over roller derby, en onze statistiek discussies. Ursula bedankt voor al je praktische hulp bij het verbeteren van mijn Engels, maar ook vooral voor het delen van lief en leed. Ik ga je missen. Ook Anita, Bart en Ruud verdienen een speciaal bedankje. Jullie staan alle drie altijd voor iedereen klaar en brengen bovendien een vrolijke vibe op de afdeling, bedankt.

Nanja, we hebben elkaar niet vaak in Delft gezien, maar gelukkig wel in Utrecht. Bedankt voor je begeleiding bij mijn eerst stapjes in de academische wereld. Onze reis naar Baden-Baden was erg fijn. Ook heb ik genoten van onze wijnavondjes en ik weet zeker dat er nog meer zullen volgen.

During my PhD traject I had the opportunity to visit the Relational Agent Group of Timothy Bickmore in Boston. I want to thank all the members of that group for the warm welcome and the great lunches. Special thanks to Ha for being my friend and the good memories we made in New York.

Ook anderen hebben me op verschillende manieren geholpen bij mijn onderzoeken. Chantal bedankt voor je mooie ontwerpen voor de app en uiteindelijk ook de omslag van deze dissertatie. Bedankt voor je doorzettingsvermogen en bereidheid om altijd mee te denken en aanpassingen te maken. Bas bedankt voor je inzet in je geduld tijdens het ontwikkelen van de Pre-treatment Motivation Module, zonder jou was dat niet gelukt. Lisanne, bedankt voor je hulp bij het beheren van de Sleepcare mailbox tijdens de Randomized Controlled Trial. Zonder jou was ik zeker omgekomen in alle e-mails. Ook wil ik alle participanten bedanken voor het meedoen aan de verschillende studies en de feedback die we gekregen hebben.

Gelukkig had ik naast het promoveren ook nog tijd voor andere activiteiten. Frans, Naomi, David, Bianca, Niels, Sheila, Qp, Susan, Alex, Lisanne, Youssh en ook de rest van de Domstad Tematen, bedankt voor de gezellige etentjes en stapavonden, waarin jullie altijd tijd maakten om naar mijn verhalen te luisteren en me zo nodig weer richting het juiste pad stuurden. Ook wil ik al mijn teamgenootjes bedanken, zowel van zaalvoetbal als softbal. Bij jullie kon ik even lekker mijn energie alsook mijn verhaal kwijt. Op nog vele sportieve jaren en wie weet een kampioenschap.

Inge, ik kan hier natuurlijk nooit beschrijven wat wij allemaal samen meemaken, maar het is altijd een avontuur. Bedankt voor je enthousiasme, je luisterend oor en je adviezen, maar bovenal voor het zijn van mijn vriendin. Ik heb genoten van al onze reizen en ik kon mijn PhD niet beter afsluiten dan een reis met jou naar Peru. Bedankt!

Als laatste, lieve mam en pap, bedankt voor jullie onvoorwaardelijke steun en liefde. Het lijkt zo vanzelfsprekend, maar dat is het niet. Jullie hebben me gemaakt tot wie ik ben en dit werk is dus grotendeels ook jullie werk. Dank jullie wel! Lieve Veronique, die onvoorwaardelijke steun en liefde krijg ik ook altijd van jou. Ik ben er super trots op dat jij mijn zusje bent. Ik hou van jullie!

Corine