Engagement with Digital Behaviour Change Interventions: Conceptualisation, Measurement and Promotion

By Olga Perski

Thesis submitted for the Degree of Doctor of Philosophy

University College London

Department of Clinical, Educational and Health Psychology

September 2018

Declaration

I, Olga Perski, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

The following work was carried out at the Department of Clinical, Educational and Health Psychology, University College London, under the supervision of Professor Susan Michie, Professor Ann Blandford and Professor Robert West. This thesis has not been submitted, in whole or in part, for any other degree, diploma or qualification at any other University.

My work was funded by a grant from Bupa under its partnership with University College London.

This thesis does not exceed the limit of 100,000 words specified by the Degree Committee.

Correspondence concerning this thesis should be addressed to Olga Perski, olga@perski.com.

Olga Perski

Signed, 24th September 2018

Acknowledgements

First of all, I would like to thank my fantastic supervisors: Susan Michie, Ann Blandford and Robert West. I have learnt so much from your expertise, detailed inputs and scientific integrity, and I am deeply grateful for your support and mentorship. I am particularly thankful to Susan for the numerous opportunities you have provided or encouraged me to seek out, which have helped me develop confidence as a scientist. I would similarly like to thank Ann for your support in navigating the complex space of interdisciplinary and mixed-methods research. I am also thankful to Robert for your enthusiasm, clear thinking and openness to new ideas. It has been a privilege to work with such inspiring leaders in the fields of behavioural science and human-computer interaction. I would also like to thank Bupa for funding my PhD and all my participants for their time and effort.

Thank you to my wonderful collaborators: Dario Baretta, David Crane, Claire Garnett, Jim Lumsden and Felix Naughton. I am deeply grateful for your intellectual and practical inputs; I have thoroughly enjoyed our discussions (both in person and through extended e-mail conversations), which have helped shape this thesis. Thank you to everyone in the Health Psychology Research Group, UCL's Tobacco and Alcohol Research Group and the UCL Institute of Digital Health; I feel lucky to have been part of such dedicated and supportive research groups. In particular, heartfelt thanks to Jamie Brown, Marta Marques and Emma Beard for your invaluable intellectual and statistical inputs. Thank you also to my 'TechSharing' buddies, particularly Nikki Newhouse, for your friendship and for sharing my enthusiasm for digital health research.

Thank you to my parents – Kerstin and Aleksander – and to my brothers – Tom, Noam and Adam – for your unwavering love and support. In particular, thank you to my mum, for teaching me to express myself through writing, and to my dad, for instilling in me a passion for health psychology and scientific research. I am also grateful to my dear friends in London and across the globe, especially Daniella, Amanda, Anton, Madde, Lali and Fabien. Finally, I would like to express my gratitude to David for your unconditional love and constant encouragement. Without you, this PhD journey would not have seemed possible.

Abstract

Digital behaviour change interventions (DBCIs) can help people change various health behaviours; however, engagement is low on average and there is a positive association of engagement with intervention effectiveness. The extent to which this relationship is confounded or subject to reverse causality is unclear, and evidence-based models of how to promote engagement are lacking. Progress is hindered by the existence of multiple definitions and measures of engagement; this hampers attempts to aggregate data in meta-analyses.

Using smartphone applications (apps) for smoking cessation and alcohol reduction as case studies, this thesis investigated how to conceptualise and measure engagement and identified factors that influence engagement with DBCIs in general, and with apps for smoking cessation and alcohol reduction in particular. Six studies using qualitative and quantitative methods were conducted. Study 1 was a systematic, interdisciplinary literature review, which synthesised existing conceptualisations and generated an integrative definition of engagement with behavioural and experiential dimensions, and a conceptual framework of factors that influence engagement with DBCIs. Studies 3 and 4 involved the development and evaluation of a self-report measure of the behavioural and experiential dimensions of engagement. Studies 2, 5 and 6 used mixed-methods to identify factors that influence engagement with apps for smoking cessation and alcohol reduction.

Engagement with DBCIs can usefully be defined in both behavioural and experiential terms: the self-report measure demonstrated promising psychometric properties and was underpinned by two distinct factors, labelled ‘Experiential Engagement’ and ‘Behavioural Engagement’. Design features that support users’ motivation to change, foster their beliefs about the perceived usefulness and relevance of the technology, and spark their interest were found to be most important in the promotion of engagement with apps for smoking cessation and alcohol reduction. These findings can be used to inform the design of new, or modification of existing, apps for these behaviours.

Impact statement

My key contribution to knowledge is the clarification of how to define and measure engagement with DBCIs. This is an important concept that can inform the development of effective DBCIs and aid the interpretation of heterogeneous trial results.

The work presented in this thesis has gained attention from researchers working across the fields of behavioural science, public health and computer science, and has generated requests for talks at Nottingham University, Gothenburg University and the International Society of Behavioural Nutrition and Physical Activity's e- & mHealth Special Interest Group. This work has led to collaborations with researchers at University College London, the University of Bristol, the University of East Anglia and the University of Milano-Bicocca.

A version of Chapter 2 has been published in *Translational Behavioral Medicine* and presented at the European Health Psychology Society's Annual Conference (2017) and the 3rd UCL Centre for Behaviour Change Digital Health Conference (2017). Details of the peer-reviewed article are as follows:

Perski, O., Blandford, A., West, R., & Michie, S. (2017). Conceptualising engagement with digital behaviour change interventions: A systematic review using principles from critical interpretive synthesis. *Translational Behavioral Medicine, 7,* 254-267. DOI: 10.1007/s13142-016-0453-1.

A version of Chapter 3 has been published in *BMC Medical Informatics and Decision Making* and presented at the Society for Research on Nicotine and Tobacco's Annual Meeting (2017) and the 3rd UCL Centre for Behaviour

Change Digital Health Conference (2017). Details of the peer-reviewed article are as follows:

Perski, O., Blandford, A., Ubhi, H. K., West, R., & Michie, S. (2017). Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. *BMC Medical Informatics and Decision Making, 17(25),* 1-14. DOI: 10.1186/s12911-017-0422-8.

A version of Chapter 4 is currently under review at *Translational Behavioral Medicine* and will be presented at the UK Society of Behavioural Medicine's Annual Meeting (2018). Details are as follows:

Perski, O., Blandford, A., Garnett, C., Crane, D., West, R., & Michie, S. (under review). A self-report measure of engagement with digital behaviour change interventions (DBCIs): Development and psychometric evaluation of the 'DBCI Engagement Scale'. *Translational Behavioral Medicine.*

A version of Chapter 6 has been published in *Digital Health* and presented at the 4th UCL Centre for Behaviour Change Digital Health Conference (2018) and the Society of Behavioral Medicine's Annual Meeting (2018). Details of the peer-reviewed article are as follows:

Perski, O., Baretta, D., Blandford, A., West, R., & Michie, S. (2018). Engagement features judged by excessive drinkers as most important to include in smartphone applications for alcohol reduction: A mixed-methods study. *Digital Health, 4,* 1-15*.* https://doi.org/10.1177/2055207618785841

My conceptual framework of factors that influence engagement with DBCIs has been applied by industry professionals (e.g. MadPow, HRW Healthcare) and policy-makers at Public Health England. The qualitative study of factors that shape smokers' and drinkers' choice of apps was covered by the magazine Science Trends, publicised to over 250,000 readers (tinyurl.com/ydelttru). I have provided industry consultancy to Smart Peak Flow Ltd, directly applying insights from this research programme to real-world problems.

Table of contents

List of tables

List of figures

List of abbreviations

App – Application

AUD – Alcohol Use Disorder

AUDIT – Alcohol Use Disorders Identification Test

AUDIT-C – Alcohol Use Disorders Identification Test-Consumption

BCIO – Behaviour Change Intervention Ontology

BCT – Behaviour Change Technique

CFA – Confirmatory Factor Analysis

CI – Confidence Interval

CIS – Critical Interpretive Synthesis

COM-B – ‘Capability’, ‘Opportunity’, ‘Motivation’ – ‘Behaviour’

DBCI – Digital Behaviour Change Intervention

EFA – Exploratory Factor Analysis

ELMP – Elaboration Likelihood Model of Persuasion

EMA – Ecological Momentary Assessments

GAMM – Generalised Additive Mixed Model

HCI – Human-Computer Interaction

HSI – Heaviness of Smoking Index

ICC – Intraclass Correlation Coefficient

IQR – Interquartile Range

IRR – Incident Rate Ratio

JITAI – Just-In-Time Adaptive Intervention

KMO – Keiser-Meier Olkin Test of Sampling Adequacy

MTSS – Motivation To Stop Scale

NHS – National Health Service

OSF – Open Science Framework

PABAK – Prevalence and Bias Adjusted Kappa

PRISMA – Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PSDM – Persuasive Systems Design Model

RCT – Randomised Controlled Trial

SDT – Self-Determination Theory

TAM – Technology Acceptance Model

UCL – University College London

UK – United Kingdom

US – United States

UX – User Experience

Contributions

Study 1 (reported in Chapter 2): I conceived of the study with Ann Blandford, Robert West and Susan Michie. I compiled the quantitative and qualitative data, conducted the narrative synthesis and wrote the chapter. All co-authors contributed to and approved the final version of the chapter.

Study 2 (reported in Chapter 3): I conceived of the study with Ann Blandford, Robert West and Susan Michie. I collected data, conducted the qualitative analysis with support from Harveen Kaur Ubhi and wrote the chapter. All co-authors contributed to and approved the final version of the chapter.

Study 3 (reported in Chapter 4): I conceived of the study with Ann Blandford, Robert West and Susan Michie. I collected data with support from Claire Garnett and David Crane, conducted the statistical analyses and wrote the chapter. All co-authors contributed to and approved the final version of the chapter.

Study 4 (reported in Chapter 5): I conceived of the study with Ann Blandford, Robert West and Susan Michie. I collected data with support from Jim Lumsden and Claire Garnett, conducted the statistical analyses and wrote the chapter. All co-authors contributed to and approved the final version of the chapter.

Study 5 (reported in Chapter 6): I conceived of the study with Dario Baretta, Ann Blandford, Robert West and Susan Michie. I collected and analysed the data with support from Dario Baretta and wrote the chapter. All co-authors contributed to and approved the final version of the chapter.

Study 6 (reported in Chapter 7): I conceived of the study with Felix Naughton, Claire Garnett, Ann Blandford, Robert West and Susan Michie. I collected the data, conducted the statistical analyses with support from Emma Beard and wrote the chapter. All co-authors contributed to and approved the final version of the chapter.

Summary of thesis

This thesis reports six studies that used a range of qualitative and quantitative methods. Study 1 (reported in Chapter 2) was a systematic review of the behavioural science and human-computer interaction (HCI) literatures, which aimed to provide an overview of different conceptualisations of engagement with digital behaviour change interventions (DBCIs), and synthesise these to develop an interdisciplinary, integrative definition of engagement. A secondary aim was to provide an overview of factors (e.g. intervention content, design features) that have been found or hypothesised to influence engagement with DBCIs. This led to the development of a two-part, integrative definition of engagement with both behavioural (e.g. amount, depth, and frequency of use) and experiential (i.e. attention, enjoyment, interest) facets, and the development of a conceptual framework which outlined factors that have been found or hypothesised to influence engagement with DBCIs.

Study 2 (reported in Chapter 3) used think aloud and interview techniques to explore what intervention content and design features are judged by potential users as important for the uptake of and engagement with apps for smoking cessation and alcohol reduction. It was found that users may select apps based on their immediate look and feel, ‘social proof’ (i.e. other users’ ratings and brand recognition) and realistic and relevant titles. Building onto the conceptual framework developed in Study 1, it was also found that intervention content and design features that enhance users’ autonomy, motivation, personal relevance and foster a sense of credibility, in addition to those that are consistent with users’ online and offline social preferences, are considered to be important for engagement with apps for smoking cessation and alcohol reduction. Study 2

was also used to gather insight into how potential users understand the term 'engagement' with DBCIs. These data were subsequently used to inform the development of a novel self-report measure of engagement with DBCIs (reported in Chapter 4).

Study 3 (reported in Chapter 4) describes the development and first psychometric evaluation of the 10-item 'DBCI Engagement Scale', which assessed the behavioural and experiential facets of the state of engagement with DBCIs. Study 3 was both practical and theoretical in scope. First, following the development of an integrative definition of engagement with DBCIs, it became apparent that no existing instrument was fit-for-purpose. The development of a novel instrument was expected to fill this gap. Second, the demonstration of adequate psychometric properties was expected to serve as an empirical validation of the proposed two-part definition of engagement, developed based on findings from Studies 1 and 2. It was expected that the state of engagement, which occurs during the momentary interaction with a DBCI, is underpinned by five dimensions: amount of use (i.e. time spent per login), depth of use (i.e. proportion of DBCI components accessed per login), attention, enjoyment and interest. The 'DBCI Engagement Scale' was evaluated in a sample of excessive drinkers who voluntarily downloaded the theory- and evidence-based *Drink Less* app, developed by researchers at University College London. Study 3 found that behavioural and experiential indicators of engagement may resolve to a single dimension, and that initial behavioural and experiential engagement did not predict future behavioural engagement. However, only a small proportion of eligible users completed the survey, which resulted in range restriction in both scale items and key outcome variables. It was therefore considered important to conduct another evaluation study in a sample with a broader range of engagement levels.

Study 4 (reported in Chapter 5) describes the second psychometric evaluation of the ‘DBCI Engagement Scale’ in a different sample of excessive drinkers who were willing to download and explore the *Drink Less* app in exchange for a financial reward, recruited through an online research platform. Study 4 found that experiential and behavioural engagement may constitute two distinct factors and that initial engagement predicted subsequent behavioural engagement. This association remained significant when adjusting for motivation to reduce alcohol consumption.

The remaining empirical studies in this thesis used novel methods to identify factors that influence engagement with alcohol reduction apps. Study 5 (reported in Chapter 6) was a mixed-methods study in which a novel ranking paradigm was used to assess what design features are considered most important for engagement with apps for alcohol reduction. In line with a ‘user-centred design’ approach, Study 5 aimed to elicit potential app users’ needs and preferences, with a view to using this information to inform the design of new, and the modification of existing, apps for alcohol reduction. Focus groups with a small number of participants and a larger, online study were conducted in parallel, addressing the same research questions. There was little agreement between participants concerning the importance of particular design features, both in the focus groups and in the online study. On average, personalisation, ‘interactive features’ and ‘control features’ were judged to be most important for inclusion in apps for alcohol reduction, as they were expected to elicit a sense of benefit and usefulness, adaptability, provide motivational support and spark users' interest. Study 5 highlighted that different features may be liked and engaged with by different users.

As studies 2 and 5 were limited by their reliance on participants' ability to predict their future preferences, experiences and behaviour, it was considered important to triangulate findings with behavioural data. Study 5 highlighted that there are individual differences in the factors that are judged to be most important for engagement with apps for alcohol reduction. Therefore, Study 6 (reported in Chapter 7) focused on the identification of within-subjects (as opposed to between-subjects) predictors of engagement. This study used a series of *N*-of-1 designs, harnessing Ecological Momentary Assessments, to examine how far within-person variability in key predictor variables identified in Studies 1, 2, and 5 (e.g. motivation to change, perceived usefulness of the app, alcohol consumption) predicted variability in the frequency (i.e. number of logins) and amount (i.e. time spent per login) of engagement with the *Drink Less* app over a 28-day period. Although different variables were found to be predictive for different users, the most consistent within-person predictor of the frequency and amount of engagement was perceived usefulness of the app.

1 CHAPTER 1 – General introduction

1.1 The promise of digital behaviour change interventions

The role of health behaviours, such as tobacco smoking and excessive alcohol consumption, in explaining morbidity and premature mortality has long been recognised [1]. Although interventions delivered face-to-face by trained healthcare professionals are both effective [2,3] and cost-effective [4], specialist services in the United Kingdom (UK) and elsewhere are facing substantial funding cuts [5] and 'brief interventions' (i.e. an intervention which takes little time to deliver and typically involves asking about smoking or drinking status) are rarely offered to patients in primary care settings [6]. Due to technological advances, Internet access and personal smartphone ownership has grown rapidly in the past decade, with 84-89% of adults in the UK and the United States (US) having access to the Internet, and 64-68% owning a smartphone in 2015-2016 [7,8]. This has led to the development of digital behaviour change interventions (DBCIs), which can be defined as "…a product or service that uses computer technology to promote behaviour change" [9]. DBCIs typically harness websites, mobile phones, smartphone applications (apps) or wearable devices to deliver behavioural support, as and when needed by the user. The potential benefits of DBCIs for patients, healthcare professionals and researchers are manifold: they can, for example, reduce the stigma associated with help-seeking in person, reach a large number of users irrespective of geographical location, be scaled up with little cost per additional user, be deeply integrated into users' daily lives and facilitate data collection in real-time [10–13].

Since the introduction of DBCIs [14], numerous randomised controlled trials (RCTs) comparing stand-alone or 'blended' DBCIs (i.e. those offering a combination of digital and face-to-face support) with wait-list or active controls have established that DBCIs are effective in helping people quit smoking, increase physical activity levels, reduce alcohol consumption, increase fruit and vegetable consumption and self-manage chronic conditions [15–23]. However, effect sizes are heterogeneous. Moderator analyses have demonstrated that some of the observed variability in effect sizes across DBCIs can be explained by differences in sample size, target population, intervention content (e.g. whether or not the digital content is theory-based, or combined with face-to-face support) and design features (e.g. aesthetics, usability) [23–26]. However, the observed variation in user engagement within and across DBCIs [27,28] may serve as an additional explanation for these heterogeneous effect sizes.

1.2 The problem of engagement

Although evidence indicates that DBCIs can help people achieve successful behaviour change, engagement with such interventions tends to be low. Eysenbach, in his 'Law of Attrition', referred to this apparent lack of engagement with DBCIs or their components as 'non-usage attrition' [27]. This phenomenon has been observed both in controlled trials of DBCIs, conducted by academic researchers, and in commercially available DBCIs, developed by industry professionals. For example, a systematic review of web-based health-related interventions found that only 50% of participants engaged with the interventions in the manner desired by the designers (i.e. interacting with all available intervention modules over a pre-specified period of time), with estimates varying between 10-90% across trials [28]. It has also been found that 25% of health and fitness apps available on the market are used only once by each user, with less than 10% of users returning to their selected app seven days after their first login session [29,30]. In light of these observations, it has been argued that there is a need to identify intervention content and design features that promote engagement with DBCIs [28].

The problem of low engagement with DBCIs is coupled with the observation of a positive association between engagement and intervention effectiveness across studies with varying characteristics (e.g. different target behaviours, delivery platforms, delivery settings) [31–34]. This has led to the hypotheses that i) engagement may be related to intervention effectiveness through a dose-response relationship, or ii) that there may be a minimum ‘effective dose’ at which users will obtain a clinically meaningful benefit from a particular DBCI (referred to in the literature as ‘effective engagement’) [35,36]. However, it is also plausible that the observed relationship between engagement and intervention effectiveness is driven by a third, unmeasured variable (e.g. greater motivation to change, better self-regulatory skills) or that it is subject to reverse causality, with users who are more successful in achieving behaviour change, being more likely to continue to engage. Given that users have self-selected into various levels of engagement in extant RCTs and observational studies of DBCIs (meaning that users have not been randomised to different ‘doses’ or levels of intensity of engagement) [37], the ability to characterise the nature of the function relating engagement and intervention effectiveness is limited at present.

In addition to the methodological problem of users self-selecting into different levels of engagement in RCTs and observational studies, the ability to aggregate results from multiple studies to identify potential moderators of engagement is limited by the presence of multiple definitions and measures of engagement. As the development of DBCIs requires knowledge not only of behavioural science, but of human-computer interaction (HCI) and computer science, different definitions and measures of engagement have emerged both within and across disciplines due to differing epistemologies and research objectives [38]. In the behavioural science literature, the predominant view is that engagement can be defined as, and measured by, intervention usage (e.g. number of logins, time spent, number of completed intervention modules) [28,36,39,40]. However, it has been noted that greater intervention usage, as indicated by more time spent on a DBCI or the completion of a greater number of intervention modules, may not necessarily reflect 'more engaged use' [41,42], suggesting that engagement intuitively comprises additional, experiential or cognitive dimensions that may not be captured solely by usage metrics. It has also been noted that spending more time on an intervention may in fact reflect slower processing speed or poor system usability, as opposed to greater intensity of engagement [42].

In the HCI literature, researchers have primarily focused on the characterisation of what it feels like to be absorbed in a digital activity, and what factors make users willing to return to a particular piece of technology or digital game. Hence, engagement has been defined in terms of the subjective experience of 'flow' or 'immersion' that emerges during the human-computer interaction, characterised by focused attention, intrinsic interest and loss of time and self-consciousness [43–45]. Consequently, engagement has typically been measured not only by usage metrics, but also through self-report questionnaires or think aloud methodology, asking about users' experiences during, or immediately after, technology use [44,46,47]. To characterise the function relating engagement with intervention effectiveness, which is needed to advance the science of DBCIs, we need a better understanding of what engagement with DBCIs is, how we can usefully measure it and what factors influence it.

1.3 Theoretical frameworks and approaches

As outlined above, multiple, inter-related scientific disciplines (e.g. behavioural science, HCI) have been concerned with the problem of engagement with DBCIs. This has resulted in contributions to separate literatures with potentially overlapping or complementary insights. An interdisciplinary perspective, which integrates existing knowledge and methodological practices from relevant disciplines, is therefore expected to help advancing our understanding of what engagement is and what factors influence it. The following section outlines key theoretical frameworks and approaches developed within the fields of behavioural science and HCI, considered relevant to the problem of engagement with DBCIs.

1.3.1 Behavioural science

As engagement with DBCIs has partly been conceptualised in behavioural terms (i.e. DBCI usage), it is expected that theoretical and practical knowledge from the behaviour change literature can inform the conceptualisation of engagement and the identification of factors that promote it. A vast number of theories about health behaviour change are currently in use: a review identified 83 different theories that have been used to predict or explain behaviour change [48]. The commonality between existing theories is that they consider psychological variables (e.g. self-efficacy, attitudes towards the behaviour) to be the most proximal predictors of behaviour change.

1.3.1.1 The COM-B Model of Behaviour

The COM-B (Capability, Opportunity, Motivation – Behaviour) model was developed with a view to synthesising the large number of existing theories of behaviour change, and posits that behaviour is part of a system of interacting components involving capability (psychological and physical), opportunity (social and physical) and motivation (automatic and reflective) [49]. Psychological capability includes the knowledge or skills necessary to perform the behaviour, and the capacity to engage in relevant memory and decision-making processes. Physical capability includes having the strength or stamina to perform the behaviour. Social opportunity refers to the affordances of one's social and cultural environment (e.g. social norms, interpersonal influences), which may act to facilitate or hinder the target behaviour. Physical opportunity includes having the time and resources necessary to perform the behaviour (e.g. having resources to pay for a gym membership). Motivation can be defined as the brain processes that energise and direct the behaviour. Reflective motivation involves conscious plans and evaluations (e.g. beliefs about consequences of the behaviour, self-efficacy), while automatic motivation includes emotional reactions, impulses and habits [49]. The development of a behaviour change intervention, irrespective of whether it is delivered face-to-face or digitally, typically involves the identification of psychological variables that maintain or hinder the target behaviour, followed by the selection of appropriate 'behaviour change techniques' (BCTs) that make up the content of the intervention [49]. These BCTs, such as goal setting or self-monitoring of the behaviour, can be combined in different ways and represent the ‘active ingredients’ of the intervention (i.e. the components of the intervention that directly affect change) [50].

1.3.1.2 The Behaviour Change Intervention Ontology

More recently, it has been acknowledged that factors beyond BCTs, such as user engagement, contribute to intervention effectiveness. Representations of how key concepts such as intervention content and delivery, engagement and intervention effectiveness inter-relate can be illustrated by means of an ‘ontology’, defined as a “…standardised representational framework providing a set of terms for the consistent description (or ‘annotation’ or ‘tagging’) of data and information across disciplinary and research community boundaries [51]. The Behaviour Change Intervention Ontology (BCIO) [9,51] proposes that engagement with a given intervention (comprising its content and the way in which that content is delivered) leads to behaviour change through influencing particular ‘mechanisms of action’, such as knowledge or self-efficacy (see Figure 1.1). These mechanisms correspond to the components of the COM-B model [51]. For example, engagement with a DBCI that includes goal-setting and feedback may lead to behaviour change through increasing the user’s psychological capability (e.g. self-regulatory skills). In addition, the BCIO predicts that delivery strategies (e.g. tailoring of content, usability, aesthetics) have a direct bearing on the extent to which users engage with a DBCI, and that the context in which the intervention is used (comprising the characteristics of the population and the setting of intervention delivery) influences user engagement. For example, demographics (e.g. gender, educational attainment) and users’ psychological states (e.g. their current level of motivation) are predicted to influence engagement with DBCIs. The BCIO also proposes that the setting in which a DBCI is delivered (e.g. the policy environment, the physical location) influences engagement with DBCIs. As the BCIO explicitly accounts for user engagement, it provides a useful means of framing the problem of engagement with DBCIs within this thesis.

Figure 1.1. The Behaviour Change Intervention Ontology (reproduced with permission from [51]).

1.3.2 Human-Computer Interaction

As engagement with DBCIs has also been conceptualised in experiential terms (i.e. the subjective experience of ‘flow’ or ‘immersion’), relevant theoretical frameworks developed within the HCI tradition are expected to complement the COM-B model and the BCIO in the study of engagement. The way in which a DBCI is delivered (e.g. elements of aesthetics, usability or tailoring) has been found to influence both engagement and intervention effectiveness [25,28], which suggests that it is important to understand the mechanisms through which such delivery strategies affect engagement.

1.3.2.1 The Technology Acceptance Model

The Technology Acceptance Model (TAM) posits that users' intentions to adopt and engage with information technology are influenced by two behavioural beliefs: beliefs about the perceived usefulness of the technology, defined as the extent to which a user believes that engagement with the system will enhance their task performance, and beliefs about the perceived ease of use of the technology, defined as the extent to which a user believes that interacting with the system will be effortless [52]. The TAM also theorises that the effect of training on how to use the technology or specific delivery strategies (e.g. aesthetics, tailoring) on usage intentions are directly mediated by perceived usefulness and perceived ease of use. There is considerable empirical support for the TAM: a meta-analysis of 59 empirical studies found that perceived usefulness is strongly related to usage intentions (r = 0.59) [53]. Perceived ease of use was also found to be significantly, albeit less strongly, associated with usage intentions (r = 0.43) [53]. However, as intentions do not always translate into action (known in the behaviour change literature as the 'intention-behaviour gap') [54], the ability of TAM to predict actual DBCI engagement is currently unclear.

1.3.2.2 The User Experience (UX) Perspective

More recently, the field of HCI has been concerned not only with the 'pragmatic qualities' of interactive products (e.g. usability), but also with their potential to accommodate 'experiential qualities' (e.g. enjoyment, beauty) and 'need satisfaction'. The concept of need satisfaction stems from Ryan and Deci's Self-Determination Theory of human motivation and wellbeing, which proposes that activities that meet the three key human needs for competence, relatedness and autonomy can enhance 'intrinsic motivation' to continue engaging in that activity (i.e. the performance of an activity for no apparent reason other than it being perceived as enjoyable or interesting in itself) [55]. This can be contrasted with 'extrinsic motivation', which refers to the performance of an activity because it is perceived as instrumental to achieving some other valued outcome. According to the UX perspective, need satisfaction refers to the ability of interactive products to satisfy users' non-instrumental needs for autonomy (defined as the feeling of being in control of one's actions), stimulation (defined as the feeling of pleasure and interest), meaning, or relatedness to other people, which can enhance users' motivation to continue engaging with the technology [56,57]. Within the UX movement, UX is defined as a dynamic, context-dependent, and highly subjective account of the human-technology interaction [57,58]. Several empirical studies have found that reports of a positive UX, characterised by beauty and need satisfaction, are positively associated with reports of perceived usability and intentions to re-engage with the technology [56,59,60]. Hence, the UX perspective is likely to complement the TAM in the study of engagement with DBCIs.

1.3.2.3 <u>The Persuasive Systems Design Model</u>

The Persuasive Systems Design Model (PSDM) argues that technology plays an important role in changing users' attitudes and behaviours (including health behaviours) through persuasion [61]. The PSDM has been widely employed in the design and evaluation of DBCIs, often with a view to promoting DBCI engagement [28,62]. The PSDM proposes four design categories that underpin persuasion: i) primary task support (e.g. reduction of complexity, tunnelling, tailoring, self-monitoring); ii) dialogue support (e.g. rewards, reminders, suggestions as to how to perform the behaviour); iii) system credibility support (e.g. trustworthiness, authority, verifiability); and iv) social support (e.g. social learning, social comparison, social facilitation). According to the PSDM, if implemented successfully, these design features will persuade users to change their attitudes or behaviours either via direct (i.e. deep) or indirect (i.e. shallow) information processing routes. In line with the Elaboration Likelihood Model of Persuasion (ELMP) [63], the PSDM argues that users' need for cognition, defined as the tendency to enjoy and seek out situations that require thinking, should moderate the effect of particular persuasion strategies on users' attitudes and behaviours. A systematic review of 83 DBCIs found that primary task support features are commonly employed in extant DBCIs, but that more extensive employment of dialogue support (e.g. reminders) is predictive of greater DBCI engagement [28]. More recently, it has been argued that persuasive design features give rise to a positive UX, which prompts users to re-engage with the technology [62]. However, empirical evidence for a link between persuasive design elements, a positive UX and increased engagement with DBCIs is lacking.

1.3.2.4 A User-Centred Design Approach

The design of health apps is often driven by the possibility of using technology, and not because the target group has expressed a need for such technology [64]. The terms 'co-design' and 'user-centred design' are used to denote design processes in which potential users influence whether, and if so, how a design takes shape [65]. The user-centred design process typically involves several, iteratively executed, stages of development, including a needs and requirements analysis, prototyping (i.e. building an early version of the software) and usability testing [66]. Although few direct comparisons of health apps designed with and without user involvement have been made (but see [67] for a meta-analysis of serious games designed with and without user involvement), user-centred design activities may help clarify the needs and preferences that have to be met for a particular digital intervention to be engaged with by the target group [64,68–70]. Approaches to identifying user needs include contextual inquiry or ethnography, which can be used to identify the key issues faced by the target group, and qualitative interviews or focus groups, which can be used to identify potential users' goals, needs and ideas for design [71]. When an initial prototype has been developed, usability testing can shed light on how the app can be refined to better meet users' needs. It was considered important to employ user-centred design approaches to identifying user needs in this thesis, as this was expected to highlight factors that promote engagement with DBCIs.

1.4 Two case studies: apps for smoking cessation and alcohol reduction

As DBCIs are available across behavioural domains and delivery platforms, with similar patterns of engagement observed across DBCIs [28,31,33], the study of the problem of engagement necessitates the selection of appropriate case studies. Tobacco smoking and excessive alcohol consumption are two of the leading causes of morbidity and premature mortality in the UK and worldwide [72]. Approximately 15% of the UK population smoke some form of tobacco (e.g. manufactured or hand-rolled cigarettes, pipe, cigars) [73]. Tobacco causes more than 6 million deaths across the world each year [72]. Alcohol consumption is more prevalent than tobacco use, with 43% of the world's adults reporting regular consumption of alcoholic beverages [74]. Excessive alcohol consumption is defined as drinking more frequently and in higher quantities than suggested by lower-risk guidelines for alcohol consumption, which typically consist of advice on weekly consumption and single episodes of drinking (often referred to as 'binge drinking') [75]. Although specific guidelines vary across countries, the UK drinking guideline states that it is safest not to drink more than 14 standard units of alcohol per week (with one standard unit containing 8 grams of pure alcohol), and that it is best to spread these units over three days or more [76]. The Alcohol Use Disorders Identification Test (AUDIT) was developed by the World Health Organisation and is a gold-standard measure for identifying individuals who drink excessively [75,77]. The AUDIT consists of ten questions about frequency of drinking, impaired control over drinking, guilt after drinking and alcohol-related injuries. Worldwide, excessive alcohol consumption causes approximately 4% of deaths [75,77], which equates to about half the number of deaths that can be attributed to tobacco smoking [78].

Pressures on national health budgets mean that face-to-face smoking and alcohol services are facing large funding cuts [5]. This has led to the development of web- and mobile phone-based interventions for smoking cessation and alcohol reduction. Similar to DBCIs for other health behaviours, meta-analyses of RCTs evaluating the effectiveness of DBCIs for smoking cessation and alcohol reduction indicate that effect sizes are heterogeneous, ranging from small to large [15,16,22,24,79]. For example, a Cochrane review of RCTs evaluating the effectiveness of DBCIs for alcohol reduction found that participants randomised to using a DBCI drank approximately 23 grams of alcohol per week less (approximately 3 UK units) than wait-list controls [16]. Moreover, a Cochrane review of RCTs evaluating the effectiveness of DBCIs for smoking cessation reported that use of an interactive DBCI, compared with a non-active control, was associated with a 15% increase in abstinence rates [15]. However, in meta-analyses comparing interactive DBCIs with active controls (e.g. face-to-face interventions or static DBCIs), or those comparing DBCIs plus human support with stand-alone DBCIs, effect sizes are heterogeneous [15,16].

More recently, smartphone apps for smoking cessation and alcohol reduction have become available. As smoking and alcohol consumption are partly driven by environmental cues that give rise to strong cravings [80,81], it has been argued that apps for smoking cessation and alcohol reduction have the potential to deliver behavioural support to users in real-time, as and when needed [82]. As smartphones are typically carried with the user throughout the day, they can be used to deliver behavioural support 'just-in-time' (i.e. pro-actively engaging users at the right time, in the right context) [83].

There are currently hundreds of smoking- and alcohol-related apps available on the market; however, only a handful of these have been designed based on theory and evidence [84–86]. For example, a content analysis of 98 popular smoking cessation apps available in the US found that only a minority of apps adhered to clinical guidelines (e.g. recommending approved medications, assisting with a quit plan) [84]. Similarly, a content analysis of 384 alcohol-related apps found that half of these were focused on entertainment, actively encouraging users to drink, as opposed to supporting users to cut down [85].

While popular smoking cessation and alcohol reduction apps vary in their specific approaches to behaviour change, commonalities in the techniques employed have been identified. For example, four independent content analyses of smoking cessation apps available in the US [84,87], UK [88] and South Korean [89] versions of the iTunes Store/Google Play Store found that at least one of the following techniques was employed in a large proportion of the reviewed apps: self-monitoring (e.g. tracking cigarettes smoked or days smoke-free), feedback on progress, advising on how to quit, rewarding abstinence, supporting identity change and hypnosis [84,88,89]. Three independent content analyses of alcohol-related apps available in the US [90], Australian, [85] and UK [86] versions of the iTunes Store/Google Play Store found that although the majority of apps actively encouraged alcohol consumption, those promoting alcohol reduction commonly employed at least one of the following techniques: self-monitoring, feedback on progress (e.g. money saved from not buying alcohol), social support (e.g. phone contact with one’s sponsor), psychoeducation (e.g. information about the negative effects of excessive alcohol use) and hypnosis (e.g. audio recordings to encourage relaxation) [85,86,90]. With regards to features aimed at promoting engagement, one review of smoking cessation apps found that some form of content tailoring was employed in 45% of apps [87] while another review identified a decline in the use of engagement features such as tailoring of content and rewards (e.g. points/badges) in smoking cessation apps between 2012 and 2014 (69.6% reducing to 45.3%) [88].

Results from the first few controlled trials of evidence-based smoking cessation and alcohol reduction apps suggest that these show promise in helping smokers to quit and excessive drinkers to cut down [82,91–96]. For example, the theory and evidence-based *Drink Less* app was designed by researchers at University College London to help adults who drink excessively to reduce their alcohol consumption through the provision of goal setting in addition to five distinct intervention modules (i.e. self-monitoring and feedback, action planning, normative feedback, identity change and cognitive bias re-training) [86,97,98]. A factorial RCT of the *Drink Less* app, which evaluated the effect of each of the five intervention components and their interactive effects on past week alcohol consumption, found that although there were no significant main effects of the intervention components, the two-way interactions between normative feedback and cognitive bias re-training, and between self-monitoring and feedback and action planning, were significant [95]. However, of 672 eligible users, only 179 (27%) completed the one-month follow-up survey. As DBCI engagement tends to be positively associated with response to follow-up (i.e. users who have disengaged with a DBCI are by definition unlikely to return to respond to follow-up measures) [27,99], this suggests that engagement with the *Drink Less* app might have been suboptimal. An RCT comparing the *REQ-Mobile* smoking cessation app with supportive text messages in a population of young adults found that the text messaging was superior to the app in achieving abstinence at three-month follow-up [91]. In addition, the frequency of engagement with the *REQ-Mobile* app was positively associated with quitting success in this sample. Hence, as initial reports of engagement patterns in smoking cessation and alcohol reduction apps appear to mirror those in DBCIs for other behaviours, they constitute two important case studies for examining the problem of engagement.

1.5 Aims of the current thesis

Using smartphone apps for smoking cessation and alcohol reduction as case studies, the aims of this thesis were:

1. To gain a better understanding of how to conceptualise engagement with DBCIs
2. To gain a better understanding of how to measure engagement with DBCIs
3. To identify factors that promote or detract from engagement with DBCIs in general, and with smoking cessation and alcohol reduction apps in particular

2 CHAPTER 2 – Conceptualising engagement with digital behaviour change interventions: A systematic review using principles from critical interpretive synthesis (Study 1)

2.1 Abstract

Background: Engagement with DBCIs is considered important for their effectiveness. Evaluating engagement is therefore a priority; however, a shared understanding of how to usefully conceptualise engagement is lacking. This review aimed to synthesise literature on engagement to identify key conceptualisations, and to develop an integrative conceptual framework involving potential direct and indirect influences on engagement and relationships between engagement and intervention effectiveness.

Methods: Four electronic databases (Ovid MEDLINE, PsycINFO, ISI Web of Knowledge, ScienceDirect) were searched in November 2015. A total of 117 articles that met the inclusion criteria were identified: studies employing experimental or non-experimental designs with adult participants explicitly or implicitly referring to engagement with DBCIs, digital games, or technology. Data were synthesised using principles from Critical Interpretive Synthesis.

Results: Engagement with DBCIs is conceptualised here in terms of both experiential and behavioural aspects. A conceptual framework is proposed in which engagement with a DBCI is influenced by the DBCI itself (content and delivery), the context (the setting in which the DBCI is used and the population using it), and the behaviour that the DBCI is targeting. The context and 'mechanisms of action' of the DBCI may moderate the influence of the DBCI

itself (i.e. content and delivery) on engagement. Engagement in turn moderates the influence of the DBCI on those mechanisms of action.

Conclusion: In the research literature, engagement with DBCIs has been conceptualised in terms of both experience and behaviour, and sits within a complex system involving the DBCI, the context of use, mechanisms of action of the DBCI, and the target behaviour.

2.2 Introduction

To date, we have not achieved a shared understanding of how to usefully conceptualise and operationalise engagement with DBCIs. This systematic review, which follows the Cochrane Collaboration’s Handbook of Systematic Reviews of Interventions [100], examines how engagement has been construed and measured in the behavioural science, computer science, and HCI literatures, and uses this to propose an integrative definition and conceptual framework of engagement with DBCIs that can be used to generate predictions and explanations of empirical observations.

The design of DBCIs requires knowledge of intervention content, delivery, interface design, and computer programming, which have traditionally been informed by separate scientific disciplines, such as behavioural science, computer science and HCI. Scientific disciplines are characterised by accumulating a body of specialist knowledge and developing a specific terminology concerned with the particular object of research [101]. Due to the multifaceted structure of DBCIs, an interdisciplinary approach, where knowledge from multiple disciplines is harnessed to develop a shared viewpoint, is required to develop a useful conceptualisation of engagement in this context [102].

As described in Chapter 1, engagement has traditionally been conceptualised differently across the behavioural science, computer science and HCI literatures, which might be due to the different epistemologies subscribed to, the differing research contexts, and the different objectives pursued. In the computer science and HCI literatures, engagement has traditionally been conceptualised as the subjective experience of ‘flow’, a mental state characterised by focused attention and enjoyment, [43]. This kind of conceptualisation might have emerged as a result of the focus on entertainment and usability of interactive technologies. In the behavioural science literature, engagement has typically been conceptualised as ‘usage’ of or ‘adherence’ to DBCIs, focusing on the temporal patterns (e.g. frequency, duration) and depth (e.g. use of specific intervention content) of DBCI use [40,103]. This kind of conceptualisation has emerged due to the observation that whilst many download and try DBCIs, sustained usage is typically low [27,30,104,105]. Henceforth, two working definitions of engagement as used in the computer science and HCI literatures (‘engagement as flow’) and the behavioural science literature (‘engagement as usage’) were used to scope the space within which this review was conducted.

Although existing systematic reviews have assessed whether particular DBCI features (e.g. tailoring, reminders) are associated with higher engagement [28,106], and whether engagement is associated with intervention effectiveness [36], it is not possible to synthesise results from these reviews or to draw any conclusions regarding the shape of the function (e.g. linear, non-linear) relating engagement with intervention outcomes due to the use of incomparable definitions of engagement [36]. In order to reduce fragmentation of research efforts, it would be useful to develop a shared understanding of how to conceptualise and operationalise engagement with DBCIs.

A conceptual framework can been defined as "…a system of concepts, assumptions, and expectations, and the presumed relationships among them" [107]. Previous conceptual frameworks of engagement have proposed multiple interacting factors (e.g. social support, sensory appeal, ease of use) that influence 'engagement as flow' or 'engagement as usage' [108–110]; however, these frameworks are either not derived from empirical observations or draw only on literature from one of many inter-related scientific disciplines. For example, the framework proposed by O'Brien and Toms [108], notwithstanding its grounding in empirical observations, drew only on research from the technology literature, and focused on 'engagement as flow' without any links to behaviour change. Conversely, the framework by Ritterband and colleagues [109] focused on 'engagement as usage', and was derived from behavioural science theory only. The model proposed by Short and colleagues [110] attempted to integrate both theoretical predictions and empirical findings from the behavioural science, persuasive design, and technology literatures, but did not do so in a systematic manner. Although the Behaviour Change Intervention Ontology proposed by West and Michie provides a starting point for organising and representing DBCIs, engagement constitutes one of many important components and is hence not examined in detail [9]. It is therefore not possible to determine whether existing frameworks of engagement sufficiently explain real-world events, or whether important aspects are missing.

The aims of this review were threefold, the second and third building on output from the first:

1. To synthesise past work on engagement, addressing the following research questions:
   a) How has engagement been defined in the selected literatures?
   b) How has engagement been measured?
   c) What factors have been found or hypothesised to influence engagement?
   d) What are the proposed relationships between engagement and intervention effectiveness?
2. To develop an integrative definition of engagement with DBCIs and specify how it can be measured.
3. To develop a conceptual framework of the direct and indirect influences on engagement with DBCIs and the proposed relationships between engagement and intervention effectiveness.

2.3 Methods

The Cochrane Handbook of Systematic Reviews of Interventions [100] and the Guidance for Undertaking Reviews in Health Care [111] were used to inform the development of the search strategy, identify inclusion criteria, select studies, and extract the data. Principles from Critical Interpretive Synthesis (CIS) were used to inform the data synthesis [112]. As CIS is one of the few methods available that affords the synthesis of qualitative and quantitative data, it was deemed to be the most suitable method. CIS is useful when a review seeks to identify a definition of a phenomenon, as it aims to produce a higher-order structure or conceptual framework (‘synthesising argument’), which is grounded in the concepts (‘synthetic constructs’) identified in the reviewed articles [112]. CIS does not propose a formal method for critically appraising the study quality and methodological rigour of included studies, but recognises that the critical evaluation and integration of disparate forms of evidence is essentially a product of the ‘authorial voice’ [113]. The evidence is critiqued on the basis of the implicit assumptions underlying the methodological decisions made in the reviewed articles. Hence, the quality of the evidence is considered in the development of the synthetic constructs, with the consideration based on the authors’ judgments. Principles of CIS have previously been employed in reviews of the health literature [114–116].

2.3.1 Criteria for considering studies for this review

All types of study designs were included except position papers. All types of information sources were included except articles that were not peer-reviewed or not available in English. Studies with adult participants (i.e. aged 18 years or older) were included, as it was expected that different factors might influence engagement in children and adult populations due to different cognitive abilities [117]. Studies specifically targeting participants with cognitive impairment or intellectual disabilities were excluded for the same reason. DBCIs and digital interventions targeting individuals with mental health or chronic physical health conditions were included as no *a priori* reason suggesting that engagement should be conceptualised differently across the included topic areas could be identified. Interventions were excluded if they did not incorporate any digital component as part of the intervention itself (i.e. face-to-face delivery only) or if the technology was used solely as a tool to deliver measurement surveys.

Studies involving recreational or educational digital games, or multimedia software (e.g. software involving animations, sound, and text) were included providing that engagement was discussed or measured. For the conceptualisation of ‘engagement as flow’, the games or technology did not need to be related to behaviour change. The primary outcome was definitions of engagement with DBCIs, digital games, or multimedia software expressed either implicitly or explicitly. Secondary outcomes included proposed direct and indirect influences on engagement, measures of engagement, and associations between engagement and intervention effectiveness expressed either implicitly or explicitly.

2.3.2 Search methods for the identification of studies

2.3.2.1 Electronic searches

A structured search of the following electronic databases was conducted in November 2015: Ovid MEDLINE (1946 – November 2015), PsycINFO (1806 – November 2015), ISI Web of Knowledge (1900 – November 2015), and ScienceDirect (1900 – November 2015). Search terms were piloted and refined to achieve a balance between sensitivity, i.e. retrieving a high proportion of relevant articles, and specificity, i.e. retrieving a low proportion of irrelevant articles [100]. An academic librarian was consulted for the validation of the databases and the final search terms. Terms were searched for in titles and abstracts as free text terms or as index terms (e.g. Medical Subject Headings) where appropriate (see Appendix 1).

2.3.2.2 <u>*Searching for other resources*</u>

Articles from adjacent fields not immediately or obviously relevant to the research questions were identified through expertise within the supervisory team [112]. The Association for Computing Machinery Digital Library (a repository for conference proceedings) and relevant journals (i.e. *Journal of Medical Internet Research, Journal of the American Medical Informatics Association, Telemedicine & e-Health*) were hand searched, and reference chaining was employed to identify additional articles of interest [100,112].

2.3.2.3 <u>*Data collection and analysis*</u>

2.3.2.3.1 Selection of studies

Articles identified through the electronic and hand searches were merged using EndNote X7 [118] to ensure consistency. Duplicate records were removed. Two researchers independently screened (i) titles, (ii) abstracts, and (iii) full texts of the identified articles against the pre-defined eligibility criteria [100]. Any disagreements were resolved through discussion, and by consulting a third researcher if necessary. Inter-rater reliability was assessed based on two coding categories (i.e. inclusion versus exclusion) after the full text screening phase with the prevalence and bias adjusted kappa (PABAK) statistic, which controls for chance agreement [119]. The following cut-offs were used: 0.40-0.59 indicates fair agreement, 0.60-0.74 indicates good agreement, and >0.75 indicates high agreement [100].

2.3.2.3.2 Data extraction and management

A pro-forma was developed to extract information about the study setting, participant characteristics, study design, data collection method, and study findings [112]. The pro-forma was piloted on a sample of included articles to ensure that relevant information was captured [100]. A second researcher independently checked the pro-forma for accuracy and completeness [111]. Due to limited resources, the data extraction was completed by one researcher.

2.3.2.3.3 Quality appraisal

CIS suggests the prioritisation of seemingly relevant articles rather than favouring particular study methodologies [120]. Judgments about the relevance and underlying assumptions of articles were made, and were incorporated into the data synthesis [112].

2.3.2.3.4 Data synthesis

Based on the principles from CIS, the data synthesis comprised the following steps:

1. Concepts identified in the full texts of included articles were labelled with codes. The research questions were used as a top-down coding frame; fragments of text explicitly or implicitly referring to definitions of engagement, measures of engagement, influences on engagement, or associations between engagement and intervention effectiveness were coded.

2. A subsample of codes were selected through random sequence generation (https://www.random.org/) for validation by a second, independent researcher to increase rigour [121]. Disagreements were discussed until consensus was reached.
3. Synthetic constructs (i.e. concepts that explain similar themes) were developed from the codes and relationships between synthetic constructs were specified.
4. The synthetic constructs and the proposed relationships between constructs were validated by a second, independent researcher. Disagreements were discussed until consensus was reached.
5. Two synthesising arguments (i.e. an integrative definition and its measurement, and a conceptual framework) were developed based on the synthetic constructs.

The synthesising arguments were refined through discussion between members of the supervisory team.

2.4 Results

2.4.1 Summary of search results

Figure 2.1 shows a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of the study selection process [122]. The electronic database search yielded 925 published articles. After removing duplicates, 560 articles remained for screening. A PABAK score of 0.88 was achieved after the full text screening phase, indicating high inter-rater reliability [100]. Due to this reliability score, the additional 31 information sources were screened by a single researcher. Of the 140 full texts screened, 117 met the inclusion criteria and were included in the data synthesis. Six qualitative studies, 27 reviews, 2 mixed methods studies, and 82 quantitative studies were included. Characteristics of the included studies are described in Appendix 2.

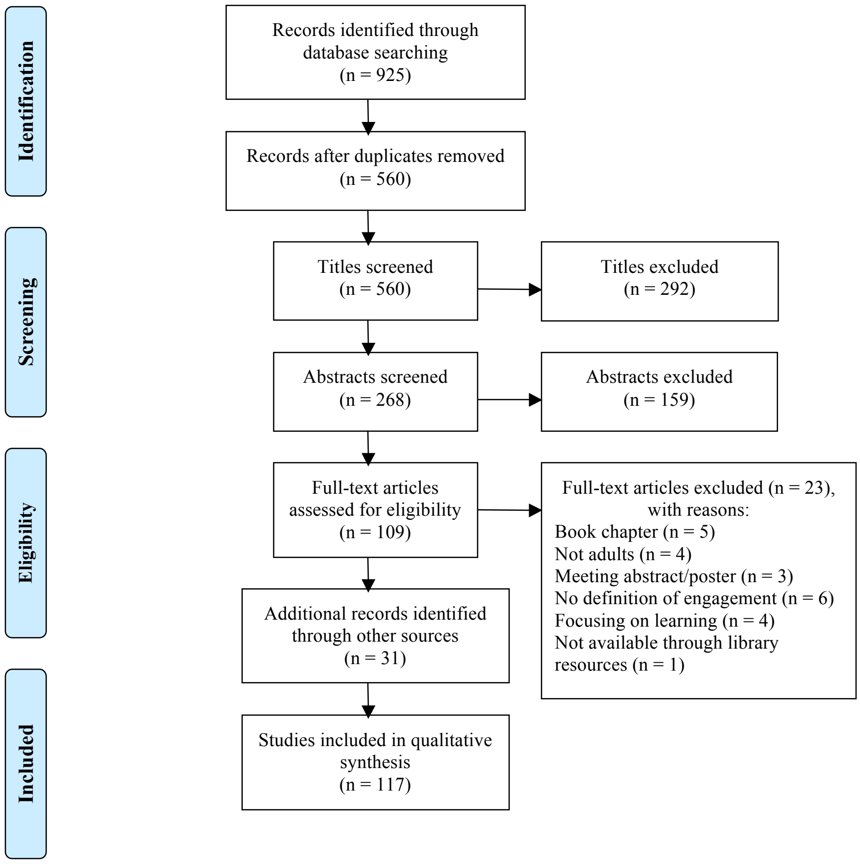


Figure 2.1. PRISMA flow diagram of the study selection process.

2.4.2 How has engagement been defined in the literature?

The following two synthetic constructs were developed: ‘engagement as subjective experience’ and ‘engagement as behaviour’.

2.4.2.1 <u>Engagement as subjective experience</u>

Engagement has been conceptualised as the *subjective experience* that emerges in the momentary interaction with a system [43,45,108]. This kind of conceptualisation was only identified in the computer science and HCI literatures. Similarities can be found between engagement and the state of 'flow', described as a mental state characterised by focused attention, intrinsic interest and enjoyment, balance between challenge and skill, and temporal dissociation (i.e. losing track of the passage of time) [43,123–126]. Similarities can also be found between engagement and the state of 'immersion' within digital gaming, characterised by cognitive absorption, the willingness to direct emotions towards an activity, and feeling cut-off from reality [44,45,127–129]. As conceptual overlap was observed between these experiential qualities, the author proposes that they can be grouped under the following cognitive and emotional states: attention, interest and affect.

2.4.2.2 <u>Engagement as behaviour</u>

The majority of articles reviewed from the behavioural science literature conceptualised engagement in *behavioural* terms, suggesting that it is identical to the usage of a DBCI or its components. Engagement has further been described as the extent of usage over time [40,130], sometimes referred to as the 'dose' obtained by participants or 'adherence' to an intervention [28,131,132], determined by assessing the following subdimensions: 'amount' or 'breadth' (i.e. the total length of each intervention contact), 'duration' (i.e. the period of time over which participants are exposed to an intervention), 'frequency' (i.e. how often contact is made with the intervention over a specified period of time), and 'depth' (i.e. variety of content used) [103,131]. In the computer science and HCI literatures, engagement has been conceptualised as the degree of involvement over a longer period of time [133], sometimes referred to as 'stickiness' [134]. A distinction has also been made between ‘active’ and ‘passive’ engagement; while the former involves contributing to the intervention through posting in an online discussion forum, the latter involves reading what others have written without commenting, also known as ‘lurking’ [135]. Engagement has also been conceptualised as a process of linked behaviours, suggesting that users move dynamically between stages of engagement, disengagement, and re-engagement [108]. As conceptual overlap was observed between these definitions, the author proposes that DBCI engagement involves different levels of usage over time.

2.4.2.3 Development of an integrative definition of engagement

An integrative definition of engagement with DBCIs was developed through the merging of overlapping conceptualisations as outlined above, in addition to the integration of the two overarching synthetic constructs. The following two-part definition is therefore proposed:

“Engagement with DBCIs is 1) the extent (e.g. amount, depth, duration, frequency) of usage, and 2) a subjective experience characterised by attention, interest and affect.”

Engagement is conceptualised as a multidimensional construct: the behavioural dimensions of engagement are underpinned by the user’s subjective experience of what it feels like to be engaged with a DBCI. Engagement is considered to be a dynamic process that is expected to vary both within and across individuals over time.

2.4.3 How has engagement been measured?

The following two synthetic constructs were developed: 'subjective measures' and 'objective measures'.

2.4.3.1 <u>Subjective measures</u>

In research settings, self-report questionnaires have frequently been used to measure engagement with digital games and DBCIs [44,46,136–144]. Qualitative approaches, such as interviews or think aloud methodology, have been employed to gain a better understanding of the nature of users' experiences of engagement with digital games and DBCIs [139,145,146].

2.4.3.2 <u>Objective measures</u>

Automatic tracking of use patterns, including the number of logins, time spent online, and the amount and type of content used during the intervention period, was the most commonly used measure of engagement in the behavioural science literature [36,40,103,106,123,147–159]. Physiological measures including cardiac activity, respiratory depth [141], and electro-dermal activity [144], and psychophysical measures such as eye-tracking [44], have been used to measure engagement in the computer science and HCI literatures.

2.4.3.3 <u>Measures relating to the integrated definition of engagement</u>

Based on the literature synthesis, the author suggests that all facets of engagement proposed in the integrative definition of engagement can in principle be measured or inferred through: 1) user-reported interaction with the

DBCI through self-report questionnaires, interview studies, or think aloud studies, 2) automated recording of DBCI use (e.g. logins, page views), and 3) recording of physiological or psychophysical correlates of DBCI interaction.

2.4.4 What factors have been found or hypothesised to influence engagement?

The following two synthetic constructs were developed: ‘context’ and ‘DBCI’. ‘Context’ was subdivided into ‘population’ and ‘setting’. ‘DBCI’ was subdivided into ‘content’ and ‘delivery’. Relationships between constructs were specified.

2.4.4.1 Context

2.4.4.1.1 Population

2.4.4.1.1.1 Psychological characteristics

Motivation was found to be positively associated with engagement across many studies, with none indicating a negative association [103,145,160–164]. As the available evidence is correlational in nature, the direction of influence cannot be established. It has been hypothesised that the relationship between motivation and engagement might be U-shaped; those who are least and most motivated to, for example, quit smoking, are hypothesised to disengage quickly from DBCIs due to failed and successful behaviour change, respectively [40].

Expectations are thought to be influential in that users are hypothesised to engage more if there is a match between their expectations and the goal of the DBCI [128,158,163,165,166]. Prior experiences of using other websites or apps, or of having tried face-to-face counselling (which may or may not have worked), might shape users' expectations of what DBCIs can provide [167].

Mental health, including low mood, anxiety, and stress, has been found to be negatively associated with engagement [32,145,158,164,168–172]. A negative association with mental health was mainly observed in studies of DBCIs targeting individuals diagnosed with a mental health condition, but was also observed in physical activity [145] and weight loss [171] interventions. Similarly, *experience of wellbeing* or believing that one does not need to work on certain issues has been found to be negatively associated with engagement [169].

Need for cognition, defined as the tendency to process large amounts of information [36,61,110,135,165]*,* and *self-efficacy* to execute a given behaviour [160,173,174] were found to be positively associated with engagement.

Personal relevance, which refers to the extent to which a DBCI is perceived to apply to the individual and their particular situation, has been hypothesised to positively influence engagement [62,146,150,175–178]. Results from interview studies indicate that participants believe that lack of personal relevance is a sufficient reason for dropping out from intervention trials [163,169,172,179].

2.4.4.1.1.2 Demographic characteristics

Age [32,39,103,135,142,145,146,148,151,155,158,168,172,174,180–184], *gender* [39,103,146,158,167,172,175,176,184], *education* [32,39,103,146,168,169,174,180,181,183,185], *employment* [168,169,181], and *ethnicity* [135,180] were found to be significantly associated with engagement. There was a trend towards a positive association between engagement and older age, higher educational attainment and being female; however, as no meta-analysis was conducted, a conclusion about the size and direction of influence cannot be drawn. *Computer literacy*, or confidence using the Internet, has been found to be positively associated with engagement [36,103,173,174,180,182,186]. However, as none of the included studies measured baseline computer skills in their designs, a firm conclusion cannot be drawn.

2.4.4.1.1.3 Physical characteristics

Physique, including baseline weight and the presence of comorbidities, was found to be negatively associated with engagement [145,155,156,168–171,180,185].

2.4.4.1.2 Setting

The *social* and *physical* environment in which a DBCI is used, has been hypothesised to influence engagement [9,109,110]. The social environment includes culture (e.g. prevailing norms), the commercial environment, media, and social cues. The physical environment includes financial resources, material resources, time pressure, physical cues, location, the healthcare system and the policy environment. *Time* [163,169,170,187] and *access* to hardware or the Internet [110,188] have been hypothesised to be positively associated with engagement.

2.4.4.2 DBCI

2.4.4.2.1 Content

DBCIs that include particular *behaviour change techniques (BCTs),* such as action plans [150], goal-setting [189], feedback [138] and self-monitoring tools [150] have been found to be associated with higher engagement [150]. *Rewards* and *incentives* have been hypothesised [106,175,176,190] or found [191] to positively influence engagement; however, evidence from trials in which the presence of rewards or incentives has been manipulated is scarce.

Social support features, referring to features that facilitate the receipt of social support, were found to positively influence engagement [148,154,192–197]. Features that decrease the feeling of loneliness or that increase motivation through competition with others include online discussion forums, gamification elements such as leaderboards that show users where they rank in a gamified system, and peer-to-peer contact [198,199]. Evidence indicates that DBCIs that provide access to such features are successful in getting users who report lower social support at baseline to engage [135,200]; however, participants who reported higher levels of social support at baseline were found to be more likely to engage with the social elements of DBCIs across a few studies [32,145,163,168].

Reminders have been hypothesised [190,201,202] or found to positively influence engagement; results from a meta-analysis indicate a positive effect of reminders on engagement [203]. However, receiving too many reminders may have a negative effect on engagement due to ‘e-mail fatigue’ [146].

2.4.4.2.2 Delivery strategies

Mode of delivery, which includes face-to-face, telephone, text message, smartphone app, website, and mass media delivery, has been hypothesised to influence engagement with DBCIs [9].

Professional support features, which include features that enable remote contact with a clinician via e-mail, telephone, or text messages, have been found to positively influence engagement with DBCIs [28,103,106,142,145,149,155,158,165,167,172,193,204–207]. However, results from an RCT of a web-based weight loss intervention in which some participants received coaching calls from a nurse indicated that participants in the coaching arm were more likely to drop out around the time of the first coaching session, suggesting a negative influence of professional support features in particular situations [155].

Control features, referring to features that make users feel that they are in control of, and are free to make choices about, how to interact with a DBCI, have been hypothesised [44,192] or found [39,130,159] to positively influence engagement. For example, results from an RCT in which participants either received content all at once or sequentially over a period of weeks suggest that participants were more likely to disengage when the content was delivered sequentially [39]. Tunnelled interventions (i.e. those that lead users through a number of predetermined steps) have been found to generate more page views compared with self-paced ones [159]. However, this may be an artefact of making users click through a pre-specified number of pages in order to progress through the DBCI.

Novelty, generated by regular content updates, has been found to positively influence engagement through preventing boredom [28,106]. However, there might be a trade-off between novelty and programme *complexity*; it has been hypothesised that participants will disengage if the intervention is perceived as too long or overly complicated [106,145,158,165,204,208,209]. It has been hypothesised that the presence of too many features may compromise a DBCI's *ease of use* [40], referring to whether or not it feels natural for the user to operate an interactive system. Ease of use has been hypothesised to positively influence engagement [156,175,210].

The *personalisation* or tailoring of content has been hypothesised [39,62,106,130,145,152,157,180,186,192,193,211] or found [40,103,136] to positively influence engagement. *Interactivity,* referring to a two-way flow of information between a DBCI and its user, has been hypothesised [108,127,129,136,150,175,212] or found [40] to positively influence engagement.

Message tone, which refers to the terminology and wording used to communicate health messages [169,176], and *narrative* [45,62,129,144,198,213], referring to the presence of a storyline, have been hypothesised to positively influence engagement. Furthermore, *challenge features* [140,175,214], *aesthetics and design* [193,212,215,216], *credibility features* [145,158], referring to features that inculcate a feeling of trust, *familiarity* [61,212,217], and the provision of *guidance* or tutorials [145,180,218] have been hypothesised to positively influence engagement with DBCIs.

2.4.5 What are the proposed relationships between engagement and the effectiveness of DBCIs?

The following four synthetic constructs were developed to explain the proposed relationships between engagement and the effectiveness of DBCIs: 'mechanisms of action', 'unmeasured third variable', 'optimal dose' and 'effective features'.

2.4.5.1 <u>Mechanisms of action</u>

Mechanisms of action proposed to mediate the effect of engagement with DBCIs on intervention effectiveness [9] include increased knowledge, motivation, affect management, cognitive restructuring, skill building [109], comprehension and practice of programme content, and increased self-efficacy [40]. A further distinction has been made between 'intervention receipt', which refers to the extent to which participants understand and can perform the skills taught, and 'enactment of intervention skills', which refers to the extent to which participants use these skills [219,220]. It has also been hypothesised that mechanisms of action, such as feeling accountable to a healthcare practitioner and relatedness to other individuals, might positively influence engagement with DBCIs [32,145,149,163].

2.4.5.2 <u>Unmeasured third variable</u>

An *unmeasured third variable,* such as higher motivation or self-efficacy at baseline, may be responsible for the observed association between increased engagement and positive DBCI outcomes. Alternatively, those who engage with DBCIs might simply be more inclined to behave healthily in general [36]. It has also been argued that the *target behaviour* itself might influence engagement [221]. For example, smokers who relapse might be more likely to stop engaging with the DBCI while those who successfully manage their cravings might be more likely to continue engaging with the DBCI.

2.4.5.3 Optimal dose

Optimal dose refers to a pre-defined level of engagement at which specific DBCIs are effective. It has been hypothesised that the receipt of an optimal dose may explain the relationship between engagement and intervention effectiveness, but that the optimal dose for particular DBCIs may vary depending on user characteristics [155,186].

2.4.5.4 Effective features

The use of specific intervention features has been found to be associated with better DBCI outcomes [155]. It has been suggested that there may be a mismatch between features that participants choose to engage with frequently and *effective features* that are causally linked to intervention outcomes [178]. For example, although users may enjoy engaging with a particular feature (e.g. filling out a food diary), thus using it frequently, infrequent use of a less gratifying feature (e.g. ‘getting support’ tools) might be more strongly associated with intervention outcomes, such as weight loss [71].

2.4.5.5 Development of a conceptual framework of engagement with DBCIs

The final aim of the review was to develop a conceptual framework specifying potential direct and indirect influences on engagement and relationships between engagement and intervention effectiveness. As the framework proposed by Ritterband and colleagues [109] and the ontology proposed by West and Michie [9] explicitly linked engagement to behaviour change, the author drew on these to structure the conceptual framework, mapping the other existing frameworks onto it. Additional factors identified in the reviewed literature not otherwise specified were also mapped onto the conceptual framework.

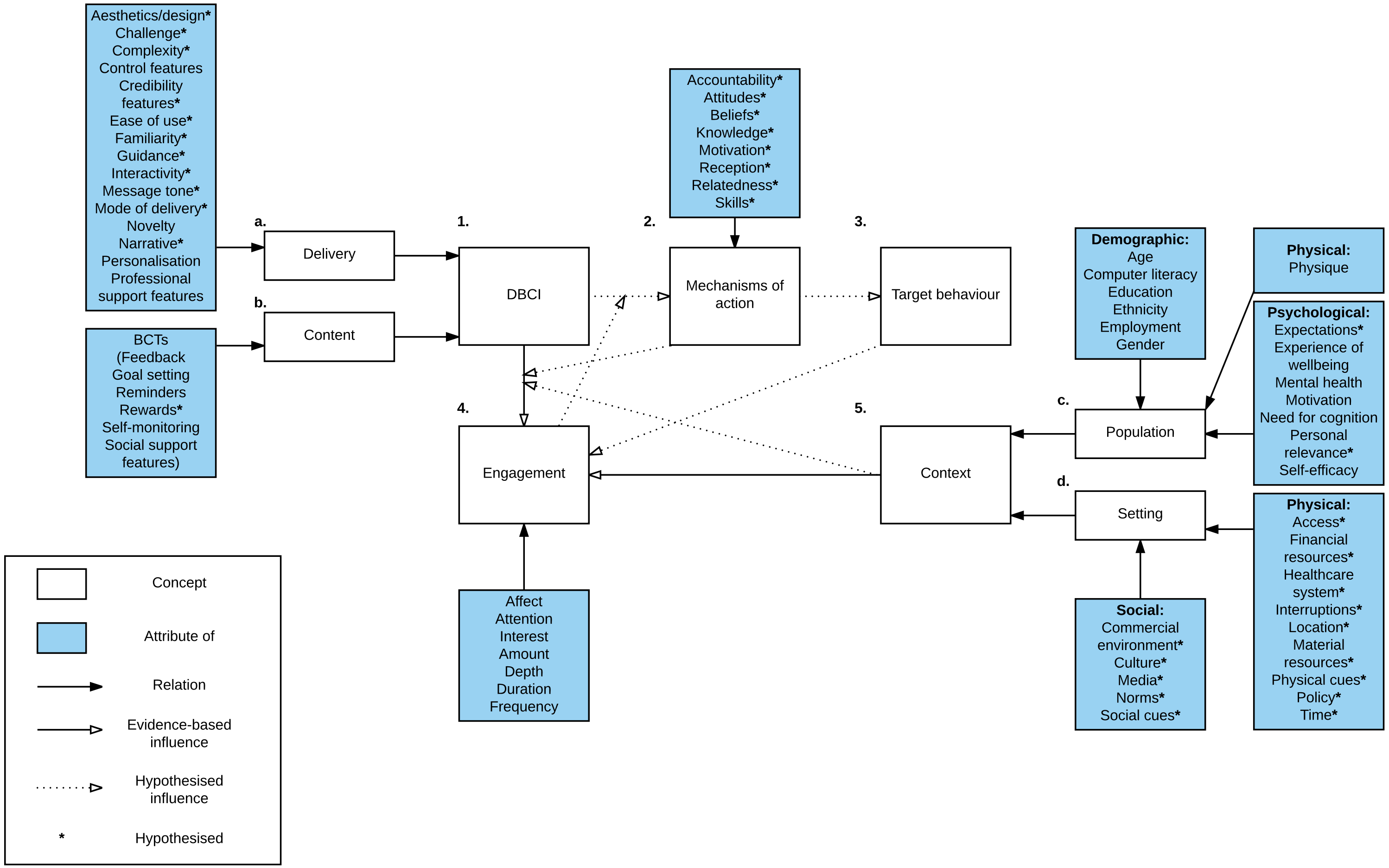


Figure 2.2. Conceptual framework of direct and indirect influences on engagement with DBCIs. Transparent boxes indicate concepts. Concepts can be defined as abstract ideas that are derived from either direct or indirect evidence [222]. Blue boxes indicate attributes of concepts. Attributes can be defined as properties that characterise a concept [223]. Solid black arrows indicate relationships between concepts and attributes. Arrows with transparent heads indicate an influence of a concept. Hypothesised influences are marked with stars.

A conceptual framework is proposed in which engagement with a DBCI influences the target behaviour through specific mechanisms of action; box 4, box 1, box 3 and box 2, respectively. Content has been found to directly influence engagement with DBCIs; box a. Delivery has been hypothesised to directly influence engagement with DBCIs; box b. The context and the target behaviour are hypothesised to directly influence engagement; box 5 and box 3, respectively. Mechanisms of action are hypothesised to indirectly influence engagement; box 2. The population (e.g. demographic, physical, and psychological characteristics) has been found to directly influence engagement with DBCIs; box c. The setting has been hypothesised to directly influence engagement; box d. Engagement is hypothesised to be indirectly influenced by the moderating influence of the context on the influence of the DBCI; box 4, box 5 and box 1, respectively. Figure 2.2 shows this schematically.

2.5 Discussion

An integrative conceptualisation of engagement with DBCIs has been developed; engagement is defined here as a multidimensional construct which can be measured through self-report questionnaires, verbal reports, automatic recording of DBCI use, or recording of psychophysical manifestations. A conceptual framework was developed, which suggests that the context of use influences engagement with DBCIs either directly or indirectly by moderating the influence of the DBCI on engagement. Mechanisms of action might indirectly influence engagement and the target behaviour might directly influence engagement with DBCIs, suggesting the presence of a positive feedback loop. The proposed relationships between engagement and intervention effectiveness are tentative, as this has not been studied extensively.

The suggested behavioural and experiential dimensions of engagement can in principle be measured or inferred in every instance of a DBCI. The content, structure, length and design of specific DBCIs tend to vary, and hence, the relevance of the different dimensions of engagement will vary accordingly. Although the intended frequency, amount, duration, and depth of use might be set to ‘1’ in a one-off intervention, the individual parameters are still present and measurable. Thus, the proposed definition of engagement allows for direct comparison across different kinds of DBCIs by including multiple dimensions of engagement at its core. This has been lacking in previous conceptualisations. Evidence of higher behavioural engagement coupled with evidence of, for example, enjoyment of using a DBCI is hypothesised to predict greater DBCI effectiveness. If this is the case, the proposed definition of engagement should provide a means of generalising findings from particular DBCIs to other, similar DBCIs. It may not be possible to evaluate the usefulness of the proposed definition prior to empirical work [224].

Although some self-report questionnaires designed to measure engagement demonstrate good validity and reliability [143,225], these typically rely on measuring engagement after, as opposed to during, the event. However, the advent of new technologies allows self-reports of engagement to be measured in real-time (e.g. through Ecological Momentary Assessments) [226]. Although physiological measures have been used to measure engagement, notably in the HCI literature, associations between physiological and self-reported measures of engagement are weak [144]. The nature of these associations thus needs to be investigated further.

Previous conceptual frameworks have been based on theoretical predictions only, or have been derived from the literature within one scientific domain [9,108–110]. In contrast, the conceptual framework proposed here is derived from theoretical predictions and empirical observations within multiple, interrelated disciplines. This endeavour was facilitated by the use of principles from CIS, which allowed the combination of a diverse set of research findings. The proposed conceptual framework of engagement is a synthesis of existing ontologies, frameworks, and models, and incorporates factors not previously included. The novel components in this framework are: ‘mental health’, ‘experience of wellbeing’, ‘familiarity’, ‘guidance’ and ‘narrative’. The negative association between poor mental health and engagement might be explained by the observation that those with poor mental health (e.g. depression) typically experience decreased self-efficacy to, for example, stop smoking or lose weight [227,228]. ‘Experience of wellbeing’ might be negatively associated with engagement due to being related to the belief that one does not need any support. ‘Familiarity’ with the design of DBCIs and ‘guidance’ might positively influence engagement because familiar examples, design conventions, or stepped how-to-use guides may inculcate feelings of comfort and ease of use. A ‘narrative’ might draw users in, increasing their interest and enjoyment. Moreover, this review identified a trend towards a positive association between engagement and older age, higher educational attainment, and being a woman, which merits further investigation. Although these demographic characteristics have been included in existing frameworks of engagement, the direction of influence has not been previously discussed. Through the use of a systematic, interdisciplinary approach, the proposed conceptual framework offers a comprehensive overview of the factors that may influence engagement with DBCIs, and hence provides a starting point for reducing the observed fragmentation of research findings.

2.5.1.1 Limitations

The lack of evidence supporting the claim that setting of use (e.g. culture, social norms, physical cues, healthcare pathway) directly influences engagement with DBCIs constitutes a limitation. This might either reflect the search terms used or indicate that this has not been investigated in the literature; a distinction between these explanations cannot be made at present. There was also a lack of evidence in support of the claim that the context of use (i.e. setting and population) may moderate the influence of the DBCI on engagement. For example, the setting of use may vary depending on the mode of delivery (e.g. computer versus mobile phone). Hence, the DBCI might indirectly influence engagement through determining the setting of use; while computers may predominantly be used at home or in a clinic, mobile phones might mainly be used on the go, which may influence the amount or depth of engagement. This hypothesis should be investigated in future research.

Another limitation is that no formal quality assessment of the included articles was conducted. However, this was in line with the chosen method, which suggests that the articles should be judged on the basis of their relevance to the research question rather than their methodological rigour. This method was selected due to the conceptual nature of the research questions. A limitation is that the data extraction and literature synthesis were conducted by a single reviewer, potentially introducing bias. Finally, the end date for the literature search (i.e. November 2015) constitutes a limitation; with the pace of technological advances and the proliferation of digital health research, relevant literature has since been published.

2.5.1.2 Conclusion

Engagement with DBCIs is conceptualised here in terms of both experience and behaviour. Engagement may be influenced by the DBCI itself, the context of use, mechanisms of action of the DBCI and the target behaviour.

2.5.1.3 Citation for the published peer-reviewed article for this study

Perski, O., Blandford, A., West, R., & Michie, S. (2017). Conceptualising engagement with digital behaviour change interventions: A systematic review using principles from critical interpretive synthesis. *Translational Behavioral Medicine, 7,* 254-267. DOI: 10.1007/s13142-016-0453-1.

See Appendix 15 for the published peer-reviewed journal article.

2.5.1.4 Next steps

To test and refine the integrative definition of engagement with DBCIs, the next steps of the thesis were to explore how potential users of apps for smoking cessation and alcohol reduction understand the term ‘engagement’ and use these insights to develop and evaluate a novel self-report measure of engagement (reported in Chapters 4 and 5). To test and extend the conceptual framework of factors that influence engagement, a qualitative exploration of smokers' and drinkers' judgments of what factors are important for the uptake of and engagement with apps for smoking cessation and alcohol reduction was conducted (reported in Chapter 3).

3 CHAPTER 3 – Smokers’ and drinkers’ choice of smartphone applications and expectations of engagement: a think aloud and interview study (Study 2)

3.1 Abstract

Background: Public health organisations such as the National Health Service in the UK and the National Institutes of Health in the US provide access to online libraries of publicly endorsed apps; however, there is little evidence that users rely on this guidance. Rather, one of the most common methods of finding new apps is to search an online store. As hundreds of smoking cessation and alcohol-related apps are currently available on the market, smokers and drinkers must actively choose which app to download prior to engaging with it. The influences on this choice are yet to be identified. This study aimed to investigate 1) design features that shape users’ choice of smoking cessation or alcohol reduction apps, and 2) design features judged to be important for engagement.

Methods: Adult smokers (*n* = 10) and drinkers (*n* = 10) interested in using an app to quit/cut down were asked to search an online store to identify and explore a smoking cessation or alcohol reduction app of their choice whilst thinking aloud. Semi-structured interview techniques allowed participants to elaborate on their statements. An interpretivist theoretical framework informed the analysis. Verbal reports were audio recorded, transcribed verbatim and analysed using inductive thematic analysis.

Results: Smokers and drinkers chose apps based on their immediate look and feel, quality as judged by others’ ratings and brand recognition (‘social proof’),

and titles judged to be realistic and relevant. Monitoring and feedback, goal setting, rewards and prompts were identified as important for engagement, fostering motivation and autonomy. Tailoring of content, a non-judgmental communication style, privacy and accuracy were viewed as important for engagement, fostering a sense of personal relevance and trust. Sharing progress on social media and the use of craving management techniques in social settings were judged not to be engaging because of concerns about others' negative reactions.

Conclusions: Choice of a smoking cessation or alcohol reduction app may be influenced by its immediate look and feel, 'social proof' and titles that appear realistic. Design features that enhance motivation, autonomy, personal relevance and credibility may be important for engagement.

3.2 Introduction

To benefit from smoking cessation and alcohol reduction apps, users must identify and select which apps to download from the myriad available on the market [84,86] and engage with them over time [35]. To the author's knowledge, no study has yet explored what factors are important in shaping users' selection and their subsequent engagement.

Although public health organisations such as the National Health Service (NHS) in the UK and the National Institutes of Health in US provide access to online libraries of publicly endorsed health apps (e.g. https://www.nhs.uk/oneyou/apps; https://www.nlm.nih.gov/mobile/) [229,230], there is little evidence to suggest that users rely on these online libraries when searching for and selecting novel apps. Rather, the two most frequently used methods of identifying new apps are to search an online store and to seek recommendations from friends and family [231]. As there are currently more than 400 smoking cessation and 700 alcohol-related apps available on the market [84,86], the onus is on the user to actively select which app to download. Notwithstanding a recent increase in the development and formal evaluation of theory- and evidence-informed apps within the research community [82,91,92,232–234], the majority of popular smoking cessation and alcohol reduction apps do not include BCTs associated with higher quitting rates in face-to-face interventions and do not adhere to public health guidelines [84–90].

As outlined in Chapter 1, due to the variable quality of available smoking cessation and alcohol reduction apps, an important goal is to determine how the design of evidence-based apps can be improved to attract users' attention in online stores and hence, increase their likelihood of being selected and engaged with [235]. The choice of any kind of app in an online store is likely to be influenced by visceral reactions to the app's design and affective responses to and cognitive processing of the app's known attributes [236–239]. Lasting positive first impressions of the visual appeal of websites are formed rapidly (within 50-500 milliseconds of exposure) and are primarily based on affective responses [236,237]. While visual appeal was highlighted by users as important when choosing from pre-specified lists of apps (e.g. health apps, games for entertainment), factors such as perceived usefulness, personal relevance, positive user ratings and prior knowledge of brand names were also considered vital [238,239]. There appears to be a lack of evidence as to how users freely choose smoking cessation and alcohol reduction apps in an online store and what factors shape their choice.

The potential benefits of apps depend not only on good choices by users but also on their subsequent engagement [35]. While evidence from RCTs indicates that features such as reminders and prompts [203], tailoring of content [40], contact with a healthcare professional [149] and simultaneous delivery of content (as opposed to sequential delivery) [39] positively influence engagement with computer- and web-delivered behaviour change interventions, little is known about the specific design features that influence engagement with smoking cessation and alcohol reduction apps.

Results from a secondary analysis of automatically recorded usage data from an RCT of a smoking cessation app indicated that users more frequently engaged with some tools compared with others (i.e. 'developing a quit plan', 'tracking smoking', 'viewing progress') [240]; however, the effect of particular design features (e.g. ease of use, tailoring of content, rewards) on engagement was not explored. In a formal consensus exercise, behaviour change and alcohol experts rated features such as ease of use, tailoring of content, feedback, aesthetic appeal and 'unique smartphone features' as likely to engage users with a novel alcohol reduction app [98]; however, it is unclear whether experts' views align with those of users from the target population. A cross-sectional survey of users' views on the functionality of an alcohol reduction app developed based on guidance from the National Institute of Clinical Excellence found that users largely held favourable views towards the app's features (e.g. an alcohol tracker, information on excessive alcohol use, notifications) [241]; however, users from the target population were not involved in the design of the app and survey respondents were not prompted to reflect on how the app's features might influence their engagement. A qualitative study that explored young adults' views on behaviour change apps and what factors contribute to their willingness to engage with such apps found that accuracy, security and immediate effects on mood were considered important for engagement while context-sensing software features and sharing on social media were considered off-putting [242]. However, no study to date has explored smokers' and drinkers' views on what design features are likely to be important for engagement with smoking cessation and alcohol reduction apps.

To better guide the selection of design features that can be included in future experimental studies (e.g. factorial RCTs), it would be useful to identify design features that smokers and drinkers judge to be important for engagement with smoking cessation and alcohol reduction apps. The present study therefore aimed to address the following two research questions through the use of qualitative methods:

1. What design features shape smokers' and drinkers' choice of smoking cessation and alcohol reduction apps?

2. What design features are judged by potential users to be important for engagement with smoking cessation and alcohol reduction apps?

This study was also used to gather insight into how potential users understand the term 'engagement' in the context of DBCIs (results reported in Chapter 4).

3.3 Methods

3.3.1 Study design

The Consolidated Criteria for Reporting Qualitative Research checklist was used in the design and reporting of this study [243]. A think aloud methodology was used to address the first research question, which involved asking participants to verbalise their thoughts, impressions and feelings whilst engaging with an app of their choice [244]. The role of the researcher in a think aloud study is to retreat to the background and only prompt participants when necessary. This method was chosen over a retrospective design due to its ability to generate real-time data on the selections made, which was considered more reliable than data generated from participants' memory. Semi-structured interview techniques were used to allow participants to elaborate on statements made during the think aloud tasks and to address the second research question. Behaviour is often influenced by unconscious processing of stimuli [245], so users may have limited insight into the factors that in fact influence their engagement with apps. However, user-centred design methods emphasise the importance of exploring users' views as part of the iterative design process in order to develop digital behaviour change interventions that accommodate the needs of the target population [246–248].

3.3.2 Theoretical framework

As the author was interested in exploring novel themes not previously identified in the literature, an interpretivist theoretical framework was used to inform data gathering and analysis [249]. Interpretivism proposes that multiple realities exist (i.e. assumes a 'subjective' rather than 'objective' reality) and that participants' accounts of their 'lived experience' are co-constructed through the interaction with and subsequent interpretations of the researcher [249,250]. Interpretivism recognises the active role of the researcher in both the elicitation and interpretation of qualitative data.

3.3.3 Participants

Smokers were eligible to take part if they i) were aged ≥ 18 years, ii) smoked cigarettes daily, iii) would consider using a smartphone app to help them stop smoking, iv) owned an Android or iOS smartphone with internet access that was capable of running apps and v) lived in or near London (UK). Drinkers were eligible to participate if they i) were aged ≥ 18 years, ii) reported an Alcohol Use Disorders Identification Test-Consumption (AUDIT-C) score ≥ 5, indicating excessive alcohol consumption [251], iii) would consider using a smartphone app to help them reduce their drinking, iv) owned an Android or iOS smartphone with internet access that was capable of running apps and v) lived in or near London (UK). Smokers and drinkers interested in using an app to stop or cut down were recruited in order to mimic real-world conditions and hence generate more valid data. It was expected that these participants would be able to more vividly imagine engaging with the apps compared with smokers and drinkers uninterested in using an app to stop or cut down [252]. For pragmatic reasons, no cut-off was imposed on cigarettes per day for including smokers in the study. As approximately 47% of English smokers are interested in using a digital intervention to stop [253], it was deemed more important to recruit smokers who were interested in using an app to stop rather than heavy or highly dependent smokers. Participants who were both smokers and drinkers were only asked about one kind of app; they were allowed to indicate a preference for what behaviour to focus on. Participants who had already tried to quit smoking/reduce their drinking using an app were not excluded. Participants who were not fluent English speakers were excluded.

3.3.4 Sampling

Participants were recruited through social media (e.g. Facebook, Twitter) and posters placed on central London university campuses. The recruitment materials stated that smokers and drinkers were invited to the laboratory to complete a few smartphone-based tasks and share their views on smoking cessation or alcohol reduction apps. Snowballing techniques were also used by asking participants to refer friends or family members interested in using an app to stop smoking or cut down on drinking [254]. Participants were recruited in batches of five until theoretical saturation was judged to have occurred (i.e. when no novel themes were identified) [255]. Preliminary data analysis was conducted after each batch of five participants to determine if more participants were needed.

3.3.5 Measures

Data were collected at baseline on: 1) age; 2) gender; 3) ethnicity, measured using the Office for National Statistics' index [256]; 4) socio-economic status, measured using the self-reported version of the National Statistics Socio-Economic Classification [257]; 5) nicotine dependence, measured using the Heaviness of Smoking Index (HSI) [258,259]; a score ≥ 4 on the HSI indicates high nicotine dependence [259]; 6) patterns of alcohol consumption, measured using the AUDIT-C [251,260,261]; an AUDIT-C score ≥ 5 indicates excessive alcohol consumption [251]; 7) motivation to stop smoking or cutting down on drinking, measured using the Motivation To Stop Scale (MTSS) [262]; 8) whether they had tried to stop/cut down in the past 12 months; 9) whether they had ever used an app to stop smoking/reduce drinking; 10) frequency of app use; 11) last time they had downloaded an app.

The MTSS is a single-item scale with seven response options: 1) “I don’t want to cut down on drinking alcohol”; 2) “I think I should cut down on drinking alcohol but I don’t really want to”; 3) “I want to cut down but haven’t thought about when”; 4) “I really want to cut down but I don’t know when I will”; 5) “I want to cut down and hope to soon”; 6) “I really want to cut down and intend to in the next 3 months”; 7) “I really want to cut down and intend to in the next month”. As the majority of available tools that tap motivation to reduce alcohol are based on the Stages of Change Model [263], for which evidence is scarce [264], the MTSS was used. Although the MTSS has yet only been validated in tobacco smokers [262], it has been successfully employed in an observational study that estimated patterns of alcohol consumption and reduction in an English sample [265].

3.3.6 Procedure

Participants read the information sheet which described the nature of the study without disclosing information that might have influenced participants’ search behaviours or verbal responses. They subsequently provided informed consent using an online screening questionnaire that assessed study eligibility and collected descriptive data (see Appendix 3). This questionnaire was hosted by Qualtrics survey software [266]. The face-to-face sessions were conducted in a private space at a London university or in participants’ homes, according to participant preference. No one else was present besides the participant and researcher except for one interview that was conducted in a space where university students were present. Interviews took place between April and June 2016. Sessions lasted between 45-75 minutes. Participants received a £20 gift voucher as compensation for their time.

3.3.6.1 Pre-task interview

A pre-session interview was held to elicit participants' expectations of apps in general and smoking cessation or alcohol reduction apps in particular (see Appendix 4). Knowledge of participants' existing beliefs about apps and their smoking/drinking identity was judged to be relevant for the interpretation of subsequent statements and reactions; for example, knowledge that a participant did not identify as an excessive drinker was subsequently used to interpret ambiguous statements or reactions towards the explored apps.

3.3.6.2 Think aloud tasks

Participants were instructed on how to think aloud (see Appendix 4) and were subsequently asked to complete a practice task: thinking aloud whilst changing the ringtone on their smartphone. Participants were then asked to complete two tasks on their smartphone. The first involved searching for smoking cessation or alcohol reduction apps in an online app store and was designed to elicit thoughts about factors that shape smokers' and drinkers' decisions to download such apps. The second task involved downloading and exploring a free smoking cessation or alcohol reduction app and was designed to gain insight into factors expected to be important for engagement (see Appendix 4). Positive reinforcement was used to ensure that participants verbalised relevant information (e.g. "You're doing well!"). When participants fell silent, prompts were used (e.g. "What are you thinking now?").

3.3.6.3 <u>Debrief interview</u>

The purpose of the debrief interview was to give participants the opportunity to elaborate on statements made during the think aloud tasks. Following the analysis of the first two batches of interview transcripts, the semi-structured interview schedule was adapted in order to elicit more data about points raised by the first 10 participants (see Appendix 4). At the end of the sessions, participants were told the full purpose of the study.

3.3.7 Data analysis

Sessions were audio-recorded, transcribed verbatim and analysed using inductive thematic analysis [267], which has previously been used to analyse data from think aloud studies involving smartphone apps [136,242]. Braun and Clarke identify six phases of thematic analysis: i) familiarising with the data, ii) generating initial codes, iii) searching for themes, iv) reviewing themes, v) defining and naming themes, and vi) producing the report [267]. Data were coded by the researcher using NVivo 10 [268] with regular discussions with members from the supervisory team. New inductive codes were labelled as they were identified during the coding process. Data were sometimes assigned to multiple codes. All codes that potentially included data relating to the study aims were recorded. The codes were reviewed one by one and findings were ordered systematically under headings. The ordered data were reviewed and revised in discussion with members from the supervisory team and were subsequently organised into themes. Theoretical saturation was judged to have occurred after 20 participants, as no new themes were identified [255]. As a quality check, a second, independent researcher reviewed the codes, themes and participant quotes. Disagreements were resolved through discussion. Agreement on the final themes was reached through discussion between members of the supervisory team. Differences between smokers and drinkers and other group differences were recorded where identified.

3.3.8 External validation

Respondent validation refers to the comparison of the researcher's interpretation of the data with participants' accounts to assess the level of agreement between the two [269,270]. A subsample of five participants (25%) was contacted and asked to review the results after the initial themes had been developed. Participants were asked to comment on whether they felt that their views were well represented and the extent to which they agreed with the interpretation of their quotes and the main claims of the narrative. Three participants returned their comments, stating that they agreed with the author's interpretations.

3.3.9 Reflexivity

Despite smoking and excessive drinking being associated with social stigma [271,272], the researcher felt that good rapport was built with the majority of participants. At the beginning of the study, the researcher asked each participant the same set of questions in the same order, but it later became apparent that a more discursive style generated more extensive data and was therefore adopted.

3.3.10 Ethical approval

UCL's Departmental Research Ethics Committee granted ethical permission (UCLIC/1213/015). Personal identifiers were removed from the data, which were stored securely, and principles of research governance were observed [273].

3.4 Results

3.4.1 Participant characteristics

Participant characteristics are reported in Table 3.1. The average age of participants was 29.7 years (*SD* = 9.2), 60% were women, 70% were of White ethnicity, 20% were of Asian ethnicity, 85% were from a high socio-economic status background and 55% of participants had made an attempt to quit smoking or cut down on their drinking in the past 12 months but had relapsed into smoking/drinking (i.e. all participants were smoking/drinking at the time of the study). Smokers had an average HSI score of 0.6 (*SD* = 1.07), indicating low nicotine dependence, and drinkers had an average AUDIT-C score of 7.0 (*SD* = 2.9), indicating excessive alcohol consumption.

Table 3.1. Participants' demographic, smoking and drinking characteristics.

ID	*Group*	*Gender*	*Age*	*MTSS**	*Made an attempt to stop/cut down in past 12 months*	*Ever used app to stop smoking or reduce drinking*	*Last time downloaded a smartphone app*	*Frequency of app use*
D1	Drinker	M	24	5	Yes	No	In the last week	Daily
D2	Drinker	M	28	2	No	No	Today or yesterday	Daily
D3	Drinker	F	28	3	Yes	No	In the last month	Daily
D4	Drinker	F	31	6	No	No	In the last month	Weekly
D5	Drinker	F	21	2	No	No	Today or yesterday	Daily
D6	Drinker	F	56	2	No	No	In the last 6 months	Monthly
D7	Drinker	F	25	2	No	No	In the last 6 months	Daily
D8	Drinker	M	24	3	Yes	No	In the last month	Daily
D9	Drinker	M	47	3	Yes	No	In the last week	Daily
D10	Drinker	M	29	5	Yes	No	In the last week	Daily
S1	Smoker	M	24	2	No	No	In the last month	Several times/week
S2	Smoker	F	25	4	Yes	No	In the last week	Daily
S3	Smoker	M	28	3	No	No	In the last week	Daily
S4	Smoker	F	20	4	Yes	Yes	Today or yesterday	Daily
S5	Smoker	F	25	5	Yes	Yes	In the last week	Daily
S6	Smoker	F	27	7	Yes	No	In the last 3 months	Daily
S7	Smoker	M	25	2	No	No	In the last month	Daily
S8	Smoker	F	45	7	Yes	No	In the last 6 months	Daily
S9	Smoker	F	33	2	No	No	In the last week	Daily
S10	Smoker	F	28	5	Yes	No	In the last 3 months	Several times/week

Note. * Motivation To Stop Scale (MTSS): 1 = I don’t want to stop smoking/cut down on drinking alcohol, 2 = I think I should stop smoking/cut down on drinking alcohol but I don’t really want to, 3 = I want to stop/cut down but haven’t thought about when, 4 = I really want to stop/cut down but I don’t know when I will, 5 = I want to stop/cut down and hope to soon, 6 = I really want to stop/cut down and intend to in the next 3 months, 7 = I really want to stop/cut down and intend to in the next month.

3.4.2 Themes

Three themes were developed in relation to the first research question and were labelled ‘immediate look and feel of the app’, ‘social proof’ and ‘realistic and relevant titles’. Five themes were developed in relation to the second research question and were labelled: ‘features that enhance motivation’, ‘features that enhance autonomy’, ‘features that enhance personal relevance’, ‘features that enhance credibility’ and ‘consistency with online and offline social preferences’. As few differences between smokers and drinkers were identified, groups were combined for the reporting of the results unless otherwise stated. A summary of the identified themes is found in Table 3.2. Supplementary quotations from the face-to-face sessions can be found in Appendix 5.

Table 3.2. Summary of identified themes.

	Theme	*Description*
1. What factors shape smokers' and drinkers' choice of apps?	*The immediate look and feel of the app*	First impressions of the app's aesthetic appeal (e.g. colour scheme, minimalist design) and usability (e.g. easy to understand, not too text-heavy).
	Social proof	The app's perceived quality, largely determined by 'social proof' (i.e. other users' ratings, recognition of credible brands/institutions).
	Realistic and relevant titles	Titles that appeared realistic and relevant to the target behaviour (e.g. "quit smoking", "reduce your drinking").
2. What factors are judged to be important for engagement?	*Features that enhance motivation*	Features that enhanced participants' motivation to stay smoke-free/reduce their drinking (e.g. monitoring and feedback, goal setting, rewards).
	Features that enhance autonomy	Features that enhanced participants' autonomy (e.g. user-controlled reminders, flexible quitting/reduction plans).
	Features that enhance personal relevance	Features that engendered a sense of personal relevance (e.g. tailoring of content, a non-judgmental communication style, gain-framed messages).
	Features that enhance credibility	Features that engendered a sense of credibility and trust (e.g. a clear privacy policy, information perceived to be accurate).
	Consistency with online and offline social preferences	Consistency with participants' attitudes towards sharing progress on social media or joining an online support community (i.e. online preferences) and their attitudes towards using the app to log cigarettes/units of alcohol or distract from cravings in social settings (i.e. offline preferences).

3.4.2.1 <u>What factors shape smokers' and drinkers' choice of apps?</u>

3.4.2.1.1 The immediate look and feel of the app

The majority of participants (14/20) stated that their choice of apps was guided by the initial appeal of icons and screenshots; however, the specific factors contributing to judgments about attractiveness differed across participants. Half of the participants (10/20) mentioned feeling drawn to apps using bright colours (e.g. light green, white), which were described as attention-grabbing or associated with health and wellbeing, while apps using dark or neon colours were considered less appealing. This divide was not universal; a few participants (2/20) felt more drawn to apps in dark colours because these were perceived as taking the quitting process more seriously.

> *Look at that! A dark screen, too many numbers. This really put me off. – D8*

When prompted to reflect on why particular designs caught their attention, many participants (9/20) mentioned that they preferred apps with minimalist or modern designs, as these were thought to signal professionalism and caring on the part of the developer, and described feeling "put off" by designs that looked "childish" or "amateurish". However, the majority of participants (11/20) were unable to articulate exactly what they liked about a particular design. This was manifested by statements about the app simply looking "nice" or having the "right" look.

> *Don't like it, yeah. I can't say more, it's just intuitive, why. It's just not something I'd particularly want to look at. - S8*

Many participants (9/20) mentioned that their choice was influenced by the app's perceived usability or simplicity, as they did not wish to invest time in apps that seemingly required too much effort, appeared to be overly complex or evoked confusion.

> *...they had these complicated graphs, and lots of information in your face, it would take you a while to read, whereas the app that I chose, it had information, it showed the progress, but it was much easier on the eye to read. - D1*

Judgments about an app's ease of use were often interwoven with judgments about its aesthetic appeal (8/20), making it difficult to single out any one factor as being more important in guiding choice.

3.4.2.1.2 Social proof

The majority of participants (15/20) mentioned that taking other people's star ratings or reviews of apps into account was vital in guiding their choice due to the lack of other guidance as to which apps are of acceptable quality. Choosing a popular app over a less popular one, determined by their respective number of downloads or list position, was thought to save time due to not having to manually filter out poor quality apps.

> *...if an app has a good rating, despite the one or two people who are not satisfied, I think it would mean that it works for the majority of people. - S1*

Many participants (8/20) mentioned feeling drawn to apps from familiar brands, organisations or developers; these were described as being more salient than other apps. When prompted to reflect on why they felt drawn to familiar brands, participants stated that they expected such apps to be of better quality than those from unknown brands; they were uninterested in information provided by developers or organisations lacking authority.

Who is [...]? Whatever, I don't care, you know. It's just some guy who came up with an app. – S6

3.4.2.1.3 Realistic and relevant titles

Many participants (9/20) mentioned that the app's title was important in guiding their choice. Titles including key words such as “quit smoking” or “reduce your drinking” were considered appealing, as these appeared to provide a realistic summary of the app's content. Participants avoided apps with titles that sounded like advertisements, such as those including the word “now”. These were thought to make empty promises about being able to help participants without providing any evidence for their statements. A few drinkers (3/10) avoided titles including the word “alcoholic”, as they did not believe that such apps would be personally relevant.

I think the title is really, really important, in terms of, don't give promises that... You've got to be really accurate and realistic, I think, to keep people interested. Don't make claims like that, just easily. – S6

3.4.2.2 ***What factors are judged to be important for engagement?***

3.4.2.2.1 Features that enhance motivation

The majority of participants (12/20) expected that regular monitoring of, for example, alcoholic beverages consumed or cigarettes smoked, and the receipt of feedback on their progress would be important for engagement. Being able to view a timeline of the days on which one had managed to stay smoke-free or drink less was expected to enhance motivation to continue, as participants did not want to “ruin their progress”.

> *That’s probably a big incentive to not smoke, because it’s just going to set that back to zero, and it’s showing you your ever increasing progress, so yeah, I do like that. - S4*

Many participants (11/20) stated that they did not expect to re-engage with apps that were too difficult to use and/or confusing. A few participants (2/20) were particularly concerned that continuously opening the app to monitor their smoking or drinking would be too effortful and hence, lead to disengagement.

Many participants (8/20) mentioned that they expected goal setting to be engaging; they believed that the achievement of a goal would make them feel good about themselves and hence, increase their motivation to achieve further goals (i.e. a positive feedback loop).

> *If you set those manageable goals, so you could achieve it, if you feel like you’re actually progressing, getting something, then you’re more likely to go back. - D10*

Of the 13 participants reacting to the provision of rewards within their selected app, approximately half (6/13) expected that the receipt of social or material rewards when achieving a goal, such as encouragement or badges, would increase their motivation to engage due to the desire to earn more rewards.

> *Doesn't [the badge] motivate you to carry on? You want to get more to prove to yourself that you can get them. – D5*

The other half of participants (7/13) was not convinced that earning virtual rewards would affect their motivation, as these participants did not attach any real value to intangible points or badges. A subtle difference between participants who had already tried to quit smoking or reduce their drinking in the past year and those who had not was observed; many (4/7) of those who had already tried to quit expressed negative attitudes towards the receipt of virtual rewards, perhaps suggesting that negative expectancy of such rewards might be linked to recent unsuccessful quit attempts.

> *I'm not really going to get any awards, am I? They're not giving me any money or presents. - D8*

3.4.2.2.2 Features that enhance autonomy

Of those expressing a desire to receive reminders to initiate engagement (11/20), the majority of these participants (9/11) wanted to control how frequently the app would contact them, as they had prior experiences of feeling bombarded or "bullied" by too many reminders.

...it was getting really, really annoying, and it bullied me a little bit too much, about me not meeting my goals that I set in the beginning when I started using it. Then it just went the other way, and it just went out the door, and I just took it off my phone. - S3

Many participants (9/20) already held firm beliefs about how to quit smoking or reduce their drinking. Smoking cessation apps that promoted a particular quitting strategy, such as quitting “cold turkey” with no option for gradual reduction, were therefore seen as inflexible. A few drinkers (4/10) expressed feeling annoyed with apps that rigidly compared their drinking patterns with the government’s recommended limits or persuaded users to have drink-free days, as they wanted to be in control of how to reduce their drinking in a meaningful way.

...it seems a bit extreme, especially when you’re not an alcoholic, why do you need a drink free day? Can’t you just have a small glass of wine with your meal? – D7

3.4.2.2.3 Features that enhance personal relevance

Tailoring of content according to individual preferences (13/20) inculcated a belief that the app was suited to the individual and that it was capable of providing effective support. For example, feedback on behavioural outcomes was estimated to be more engaging if it was tailored to the individual’s needs and preferences.

I'm supposed to be motivated by how much money I've saved. That doesn't make sense to me. I think I should be motivated by how my health might have improved. I don't like this app. It's not going to help me. - D6

Information perceived as "preachy" or patronising made participants feel judged or nagged (9/20). This resulted in refusals to take the information seriously due to the desire to rebel against advice on what one "should" do.

I think I'm more likely to listen to practical advice rather than finger wagging… - S9

Some participants (6/20) mentioned that they wanted information about the positive effects of quitting or cutting down (i.e. 'gain-framed' messages). Information about health consequences that focused on the negative aspects of past smoking or drinking (i.e. 'loss-framed' messages) made participants (7/20) feel disempowered due to the inability to change past actions. Information focusing on the negative consequences of future smoking made some participants feel indifferent due to the inability to imagine one's future self.

Great. I started smoking when I was 13 and back then, I was smoking 40 cigarettes a day. - S3

A few drinkers (3/10) were sensitive to terminology perceived as "serious" or harsh, especially when terms such as "alcoholic" or "addict" were used. They were quick to distance themselves from apps using such terminology, as they appeared to assume that these must be catered to individuals who, unlike them, were dependent on alcohol. Smokers were more accepting of the use of the term “addict”.

> *“Add an addiction.” OK, quite serious... Wow! “I’ve been clean for...” That’s some serious terminology. - D10*

3.4.2.2.4 Features that enhance credibility

Many participants (8/20) mentioned that they felt uneasy about having to create an account with their personal e-mail address or allow access to the phone’s location services in order to use their selected apps, as they were worried that their information would be passed on to third parties.

> *One thing is that I tend to not like apps that require so much data about my location services, because, I don’t know, but obviously they sell on apps, so I think I’m quite wary of telling people too much about my data... - S10*

However, a few participants (3/20) mentioned that their concerns were mitigated if a message about the app’s policy on privacy and confidentiality was provided due to feelings of trust. A few participants (2/20) explicitly stated that they had no concerns regarding privacy in the context of apps.

> *It then says: “Your data will be anonymised and not shared with anybody other than for our research”, which is nice to tell people for confidentiality reasons. - D7*

Information judged to be inaccurate was met with scepticism by many participants (8/20) as errors and inconsistencies were thought to undermine the app's credibility. Participants did not want to waste time on inaccurate advice, as this was deemed to be untrustworthy.

> *I think it's really important that these sorts of sites and apps have the most current, up-to-date information, in order to get me to trust them, and take on board what they're telling me. - D2*

3.4.2.2.5 Consistency with online and offline preferences

Of the participants who reacted to the provision of social support features within their selected apps (10/20), such as sharing progress on social media (e.g. Facebook, Twitter) or joining an online community, few (4/10) expressed a desire to engage with such features; smoking and drinking were seen as private behaviours that are unacceptable to share with one's wider social network. Participants anticipated that sharing such information with others would generate pity rather than support.

> *...what do I want to get from that? I'm not going to get endorsements, I'm just going to get a few sad likes that are going to be quite patronising to me… - S3*

A subtle difference was observed between those who had tried to quit smoking or reduce their drinking in the past year and those who had not; the former appeared to judge social sharing to not be engaging due to the anticipation of added pressure rather than increased support while the latter expressed more favourable attitudes towards social support features, especially those enabling users to join an online support community. Participants who had not made an attempt to quit expected that connecting with others in a similar situation might help stick to one's goals due to increased motivation.

Beliefs about the capability of apps to provide timely support when experiencing a craving were mixed. Many participants (7/20) struggled to see ways in which engagement with an app would influence their waning resolve. A few smokers (3/10) believed that doing a breathing exercise to assuage cravings would be helpful in the moment, but they did not want to use distraction games when socialising with others, who might find this behaviour strange.

> *Obviously, if you're in a bar, you're not going to be like: "I'm sorry guys, I just need to play my game." Maybe when you're home alone, it could be useful. – S5*

When imagining logging drinks consumed in social situations, a few drinkers (2/10) mentioned that they anticipated feeling embarrassed or uncomfortable, as others might find such behaviour "odd" or "rude" and hence, stop inviting them to the pub.

> *If I pull it out and start pressing it every time I've had a drink, they're going to start thinking that I'm odder than I really am. – D9*

3.5 Discussion

This study found that the immediate look and feel of apps, social proof and realistic and relevant titles shape smokers' and drinkers' choice of apps. Features that enhance motivation, including monitoring and feedback, goal setting, ease of use and rewards, and those that enhance autonomy, including flexible prompts and quitting strategies, were judged to be important for engagement. Participants also expected that features that engender a sense of personal relevance, such as tailoring of content according to individual preferences and the use of a non-judgmental communication style, and those that engender a sense of credibility, including privacy and accuracy, would be engaging. Moreover, consistency with one's online and offline social preferences was considered important for engagement. Few differences were found between smokers and drinkers.

The finding that the immediate look and feel of apps influenced participants' choice is consistent with the argument that visceral reactions to an app's design generate lasting positive first impressions [236,237]. However, other people's app ratings and the perceived relevance of titles were also considered important. This supports the suggestion that both affective responses and cognitive processing of an app's attributes influence users' choice of apps [238,239].

These results are consistent with a number of well-established findings. Firstly, the finding that prompts, rewards, ease of use and tailoring of content according to individual differences were expected to be important for engagement supports previous research into computer-delivered smoking cessation and alcohol reduction interventions [39,176,190], results from content analyses of smoking cessation apps [87,88] and findings from a formal expert consensus study [98]. Secondly, the finding that the app's communication style was judged to be important for engagement is consistent with previous research suggesting that the 'tone of voice' of digital behaviour change interventions may evoke strong negative emotions and hence, cause participants to disengage [169]. Moreover, the finding that privacy and accuracy are expected to be important for engagement due to feelings of trust replicates research into other kinds of digital behaviour change interventions [116,145,158].

A frequently mentioned justification for using smartphone apps to deliver complex behaviour change interventions is that these are capable of delivering support as and when required, or 'just-in-time' [274,275]. As participants in the present study expressed concerns about engaging with smoking cessation and alcohol reduction apps in social settings due to anticipated embarrassment, this adds nuances to the assumption that smokers and drinkers want timely behavioural support irrespective of context. A recent study that employed geofencing (i.e. a software feature that uses the phone's global positioning system to set up geographical boundaries) to deliver context-aware smoking cessation support found that only a small proportion of pre-quit smoking reports (6.1%) were logged in social situations [276]. One of the reasons for this, as evidenced in follow-up interviews with participants, was fear of appearing rude to other people. This finding is also consistent with views expressed by young adults in a qualitative study exploring opportunities and challenges for behaviour change apps, who questioned the accuracy of context-sensing features [242].

Consistent with previous findings [242], smokers and drinkers in the present study did not want to share progress with their wider social networks due to the belief that others would pity rather than encourage this. It has been found that so-called ‘closet’ quit attempts (i.e. attempts to stop smoking without disclosure to anyone) are common among smokers [277]. As non-disclosure does not appear to be associated with a decreased likelihood of cessation success [277], this may be interpreted to suggest that social sharing should not be considered a ‘one-size-fits-all’ approach.

Care should be taken not to overstate the importance of the present findings due to the subtle group differences observed and the small sample size. However, it was found that attitudes towards joining an online support community and attitudes towards the receipt of virtual rewards appeared to differ depending on whether participants had made an attempt to quit/cut down in the past year. This suggests that individuals may differ in the factors that influence their judgments of engagement features. Future research should explore whether individuals may respond differently to social support features and rewards depending on their demographic and/or psychological characteristics.

3.5.1 Limitations

The method chosen to elicit data involved asking participants about their expectations about what factors would be engaging. As evidence suggests that the magnitude of relationships between beliefs and attitudes, intentions and actual behaviour are modest [54], further research is required to assess whether the inclusion of the features judged by participants to be important for engagement in the present study is in fact accompanied by higher levels of engagement. Although reliable methods for determining the potential of health apps to engage users (e.g. the Mobile Application Rating Scale [278]; a coding scheme developed by Ubhi and colleagues [279]) are available, the predictive validity of such scales (i.e. the scales' ability to predict actual levels of engagement) has not been evaluated. As the purpose of the present study was to explore smokers' and drinkers' views of apps, consistent with a user-centred approach to intervention design [246–248], think aloud methodology and semi-structured interview techniques were deemed to be more appropriate than existing quality scales. It has been argued that the use of think aloud methodology to elicit data might be problematic as it is cognitively demanding for participants to complete the assigned tasks whilst verbalising their thoughts [280]. However, this issue was mitigated by conducting debriefing interviews to allow participants to elaborate on their statements.

The boundary between aesthetic appeal and perceived usability was often unclear in participants' explanations, highlighting the difficulty in articulating precisely why particular designs are considered more attractive than others and hence, indicating that the data generated here might be imprecise. However, ratings of beauty have been found to be strongly associated with ratings of perceived usability in other settings [59]. This emphasises the complexity of trying to dissociate these constructs and suggests that these findings are consistent with the published literature [236,237]. Additional insight into how smokers and drinkers select apps (e.g. specific search terms used, non-conscious selection processes) might be gained from screen recordings or the use of eye tracking methodology.

As participants in the present study were predominantly of White ethnicity from high socio-economic status backgrounds and smokers indicated low levels of nicotine dependence it is possible that these findings do not generalise across the target population. However, participants reported similar levels of motivation to stop compared with a large, representative sample of English smokers (*N* = 2,483): 35% in the present study versus 39% of English smokers in the earlier study indicated a MTSS score of ≥ 5 [262]. The finding that few smokers and none of the drinkers in the present study had ever used an app to quit smoking/reduce their alcohol consumption may be interpreted to suggest that the real concern is not how users decide which app to use, but rather, that it is more important to gain insight into what makes smokers and drinkers decide to use an app in the first place. Little is known about the uptake of smoking cessation and alcohol reduction apps in the general population of smokers and drinkers; however, findings from an ongoing series of cross-sectional household surveys of representative samples of the English population indicate that although half of smokers expressed an interest in using digital smoking cessation interventions (e.g. websites, smartphone apps), fewer than 1% had in fact used such interventions to support a quit attempt in the past year [253]. Hence, an alternative interpretation is that, according to available statistics, the present sample appears similar to the target population with regards to previous app use.

3.5.2 Conclusion

Smokers and drinkers interested in quitting or cutting down using a smartphone app choose apps based on their immediate look and feel, social proof and titles judged to be realistic and relevant. Features that enhance motivation, autonomy, personal relevance and credibility, and those that are consistent with users' online and offline social preferences are rated by participants as important for engagement.

3.5.3 Citation for the published peer-reviewed article for this study

Perski, O., Blandford, A., Ubhi, H. K., West, R., & Michie, S. (2017). Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. *BMC Medical Informatics and Decision Making, 17(25),* 1-14. DOI: 10.1186/s12911-017-0422-8.

See Appendix 15 for the published peer-reviewed journal article.

3.5.4 Next steps

The next steps of the thesis were to develop and evaluate a self-report measure that taps the experiential and behavioural dimensions of engagement with DBCIs (reported in Chapters 4 and 5).

4 CHAPTER 4 – A self-report measure of engagement with digital behaviour change interventions (DBCIs): Development and psychometric evaluation of the 'DBCI Engagement Scale' (Study 3)

4.1 Abstract

Background: Engagement with DBCIs is a potentially important mediator of effectiveness; however, we lack validated measures of engagement. This study describes: 1) the development of a self-report scale that captures the behavioural and experiential facets of engagement; and 2) the evaluation of its psychometric properties in a real-world setting.

Methods: A deductive approach to item generation was taken. The study sample consisted of adults in the UK who drink excessively, downloaded the freely available *Drink Less* app with the intention to reduce alcohol consumption, and completed the scale immediately after their first login. Five types of validity (i.e. construct, criterion, predictive, incremental, divergent) were examined using Exploratory Factor Analysis (EFA), correlational analyses, and through regressing the number of subsequent logins in the next 14 days onto total scale scores. Cronbach's α was calculated to assess internal reliability.

Results: A 10-item scale assessing amount and depth of use, interest, enjoyment and attention was generated. Of 5,460 eligible users, only 203 (3.7%) users completed the scale. Seven items were retained, and the scale was found to be unifactorial and internally reliable (α = .77). Divergent and criterion validity were not established. Scale scores did not predict the number of subsequent logins (B = .02, 95% CI = -.01, .05, p = .14).

Conclusions: Behavioural and experiential indicators of engagement with DBCIs may constitute a single dimension, but low response rates to engagement surveys embedded in DBCIs may make their use impracticable in real-world settings.

4.2 Introduction

Although many different measures of engagement are currently in use, including self-report scales and objectively recorded usage data [35,281], an instrument that captures both the behavioural and experiential facets of engagement is lacking. For example, although the *User Engagement Scale* [46], the *eHealth Engagement Scale* [143], the *Flow State Scale* [47], the *Immersion Experience Questionnaire* [44], the *Personal Involvement Inventory* [282] and the *Mobile Application Rating Scale* [278] capture a range of experiential facets (e.g. stimulation, enjoyment), they do not consider the behavioural facets of engagement (see Table 4.1 for an overview of extant self-report scales). Automatically recorded usage data have typically been employed as a behavioural index of engagement [39,40,103,283], but it is unclear whether such records provide a valid measure of the experiential facets of engagement (e.g. attention, interest). A validated measure of engagement that could be used by researchers, healthcare practitioners and industry professionals, irrespective of having access to the DBCI's raw data, would be practically useful. Therefore, the present study aimed to develop and validate a new self-report scale that captures both the behavioural and experiential facets of engagement.

Table 4.1. Overview of psychometric properties of existing self-report measures of engagement with DBCIs.

Self-report scale	Description	Construct validity	Reliability	Criterion validity	Divergent validity	Predictive validity
User Engagement Scale [46]	A 123-item scale, designed to measure the following 10 sub-dimensions of engagement: 'aesthetics', 'affect', 'focused attention', 'challenge', 'control', 'feedback', 'interest', 'motivation', 'novelty' and 'perceived time'.	Exploratory Factor Analysis (EFA) indicated a six-factor solution: 'focused attention', 'perceived usability', 'aesthetics', 'endurability', 'novelty', and 'felt involvement'. The factor solution did not replicate in a new sample [284]. The authors recognised that their definition of engagement contains attributes that predict, rather than are part of, the focal construct.	Cronbach's α was calculated to assess internal consistency reliability for each factor, ranging from .72-.90.	N/A	N/A	N/A
eHealth Engagement Scale [143]	A 12-item scale, designed to assess the following sub-dimensions of engagement with digital health information: 'absorbing', 'attention-grabbing', 'stimulating', 'surprising', 'suspenseful', 'thought-provoking', 'clever', 'convincing', 'balanced', 'believable', 'dull' and 'hip/cool'.	Confirmatory Factor Analysis (CFA) indicated acceptable fit of a four-factor model: 'involving', 'credible', 'dull' and 'hip/cool'.	Cronbach's α was calculated to assess internal consistency reliability for each factor but is not reported.	N/A	N/A	Assessed the scale's ability to predict aggregate scores on three proximal outcomes (e.g. "The information made me feel more confident that I can do something"). The four-factor solution accounted for 56% of variance in the proximal outcome.

Table 4.1. *Continued.*

Self-report scale	**Description**	**Construct validity**	**Reliability**	**Criterion validity**	**Divergent validity**	**Predictive validity**
Flow State Scale [47]	A 54-item scale, designed to measure the following 9 sub-dimensions of the state of 'flow' [43]: 'challenge-skill', 'action-awareness', 'clear goals', 'unambiguous feedback', 'concentration', 'sense of control', 'loss of self-consciousness', 'transformation of time' and 'autoletic experience'.	A series of CFAs, resulting in the removal of 18 items, indicated that the a priori nine-factor structure was supported.	Cronbach's α was calculated to assess the internal consistency reliability for each factor, ranging from .80-.86.	N/A	N/A	N/A
Immersion Experience Questionnaire [44]	A 33-item scale, designed to measure 8 sub-dimensions of the state of 'immersion' during digital game-play: 'temporal dissociation', 'focused immersion', 'heightened enjoyment', 'control and autonomy', 'curiosity', 'emotional involvement', 'transportation to a different place' and 'attention'.	EFA indicated a five-factor solution: 'cognitive involvement', 'real world dissociation', 'challenge', 'emotional involvement' and 'control'.	N/A	N/A	N/A	N/A

Table 4.1. *Continued.*

Self-report scale	Description	Construct validity	Reliability	Criterion validity	Divergent validity	Predictive validity
Personal Involvement Inventory [282,285]	30-item scale, designed to measure the ‘motivational state of involvement’ with different commercial products [282], measured using bipolar adjectives.	EFA, after removing 10 items, indicated a one-factor solution.	Test-retest reliability indicated that item-to-item correlations between Time 1 and Time 2 (3 weeks later) ranged from .31-.93.	Scale scores for products (e.g. car, jeans) were found to correspond to previous classifications of such products into low or high involvement categories.	N/A	N/A
Mobile Application Rating Scale [278]	23-item scale, designed to function as a quality assessment tool for mobile health apps, assessing the following 4 sub-dimensions: ‘engagement’, ‘functionality’, ‘aesthetics’ and ‘information quality’.	N/A	Inter-rater reliability, calculated using the intra-class correlation coefficient, ranged from .5-.83. Internal consistency reliability, calculated using Cronbach’s α, ranged from .80-.93.	N/A	N/A	N/A

4.3 Methods

4.3.1 Scale development

4.3.1.1 Construct development

The construct of interest was developed through three iterative steps [286]: i) defining the conceptual domain to which the construct belongs (e.g. thought, feeling, behaviour, outcome); ii) defining the entity to which the construct applies (e.g. person, task, process, relationship) and how stable it is expected to be over time, across situations and across cases; and iii) defining the set of fundamental attributes or characteristics that are necessary and sufficient for something to be an instance of the construct. Two data sources were drawn upon to generate a definition of the construct: a systematic review of the behavioural science, computer science and HCI literatures (reported in Chapter 2) and an empirical think aloud and interview study with potential users of smartphone apps for smoking cessation and alcohol reduction (study methodology reported Chapter 3).

4.3.1.1.1 Conceptual domain

Existing definitions of engagement identified in Chapter 2 could broadly be categorised into one of two conceptual domains: 'engagement as subjective experience', incorporating emotional and cognitive facets of engagement, and 'engagement as behaviour' [281]. A similar distinction was made by participants in the think aloud and interview study, who described feelings of attention, enjoyment and interest when engaging with a DBCI (see Table 4.2). Participants also highlighted the behavioural facets of engagement: they described engagement as frequent DBCI use over time, spending time on the DBCI when deciding to use it, and interacting with many, rather than a few, of the DBCI's features (i.e. frequency, amount of use and depth of use). It was therefore hypothesised that engagement spans two conceptual domains: an experiential domain (with cognitive and emotional facets) and a behavioural domain.

Table 4.2. Summary of themes pertaining to participants' understanding of the term 'engagement'.

Theme	Description	Example quotations
Attention	Participants described feeling ‘drawn in’ or ‘sucked in’ by engaging apps. Engagement was thought to involve sustained attention on, and active involvement with, the app’s interface and content.	“An engaging app, it’s something that draws you in…” – D5
		“Engagement is how it would hold your attention, whether it does that successfully.” – S9
Enjoyment	Engagement was thought to involve feelings of enjoyment and fun.	“I suppose how positively I think about it. If someone were to ask me: “Do you enjoy using it,” and I said: “Yes, 8 out of 10,” then that would be a form of engagement.” – D8
Interest	Engagement was thought to involve feelings of interest and stimulation. Boredom was seen as the opposite of app engagement.	“…if it interested me, stimulated me…” – S4
		“…after a few minutes on this app, I’m actually kind of bored already.” – S3
Extent of DBCI use	Engagement was described in behavioural terms; participants thought that engagement comprised the frequency, amount and depth of use.	“If you’re more engaged with an app, you’re going to use it daily, or more frequently…” – D7
		“…how much time I spend on it on each session of use…” – D8
		“How likely you are to use it for its full purpose.” – D9

4.3.1.1.2 Type of entity

As engagement has been found to vary within users over time and across DBCIs, often as a function of person- or technology-specific attributes (e.g.

motivation to change, self-efficacy, tailoring, aesthetics) [281,287–289], it was hypothesised that engagement can usefully be conceived of as a state rather than a trait [290]. The two data sources did not help clarifying whether the state of engagement is best conceived of as a task-specific construct (i.e. applicable only in situations where a DBCI is present) or whether it extends to other objects. For the purpose of the present study, it was hypothesised that the state of engagement is task-specific, as this implies that it is sufficient to consider situations in which the object of interest is a DBCI, rather than any other material object.

4.3.1.1.3 Necessary and sufficient conditions

Two behavioural indicators and three experiential indicators were identified as particularly important for determining the intensity of the state of engagement: amount of use, depth of use, attention, interest and enjoyment. First, spending time on a DBCI (i.e. 'amount of use') and accessing at least one of its components (i.e. 'depth of use') were both considered necessary for engagement. Spending time on a DBCI, though not actively using it (e.g. posting in an online forum), was considered necessary as research shows that 'lurking' in online discussion forums (i.e. reading others' comments without actively contributing) can help people achieve behaviour change [37]. As research shows that unique behaviour change techniques are independently associated with successful behaviour change [291–293], the range of components accessed was considered necessary to determine the intensity of DBCI engagement. The behavioural indicators were hypothesised to be jointly insufficient for someone to be engaged, as a user may scroll through information on an app without paying attention to its content. Therefore, three experiential indicators were also considered necessary for engagement: paying attention to the DBCI ('attention'), feeling interested in it ('interest') and experiencing enjoyment whilst using it ('enjoyment'). It is widely accepted that the process of selective attention helps allocating limited resources to specific stimuli, and that the function of interest is to direct attention towards important stimuli [294–297]. Although the two data sources also indicated that enjoyment is a key aspect of engagement [281], it is unclear whether this is a necessary condition for someone to be engaged, as it may be possible to pay attention to an app and be interested in its content without necessarily experiencing enjoyment. Given the lack of evidence at present, it was hypothesised that the two behavioural and three experiential indicators were necessary and jointly sufficient for engagement.

4.3.1.2 <u>Item generation</u>

A deductive approach to item generation was taken, meaning that the theoretical definition of the construct is used as a guide to generate scale items [298]. An initial pool of 18 items was generated by the researcher based on the theoretical definitions of the five indicators of engagement (i.e. 'amount of use', 'depth of use', 'attention', 'interest', 'enjoyment'). To mimic everyday language, items were designed to capture the intensity of the relevant thoughts, feelings and behaviours (e.g. "How strongly did you experience enjoyment?"; "How much time do you roughly think that you spent on the app?"). Agreement on the set of initial items was reached through discussion between members of the supervisory team. Although some of the resulting items resemble those from existing scales (reviewed in Table 4.1), the researcher did not explicitly draw on these. The focus was to develop items that demonstrate theoretical coherence, as opposed to novelty. Two items representing the researcher's best bets for a short measure of engagement were also developed (i.e. "How engaging was the app?"; "How much did you like the app?").

4.3.1.3 Item scaling

As the questionnaire was designed to be administered online and accessed through platforms with potentially small screens (e.g. smartphones), 7-point scaling was used where possible, with higher scores indicating greater intensity of engagement. Scale end- and mid-points were anchored to contextualise the response options: 'not at all'; 'moderately'; 'extremely' [299].

4.3.1.4 Content validity

Following the methodology in [300] and [301], a group of 10 behavioural scientists and 10 human-computer interaction experts were recruited from the author's networks (i.e. 'experts') and a group of 50 adult respondents recruited through Amazon's Mechanical Turk (i.e. 'non-experts') were invited to complete a 'content adequacy task' to determine the scale's content validity.
Respondents were asked to classify the randomly ordered items into one of six categories (i.e. 'amount of use', 'depth of use', 'interest', 'attention', 'enjoyment', plus an 'unclassified' category). The task was hosted on Qualtrics [266] and was completed remotely without any researcher input. A minimum of 70% of respondents had to correctly classify an item for it to be retained [300,301].

Of the 18 initial items, two items tapping 'interest', three items tapping 'attention', five items tapping 'enjoyment' and one item tapping 'amount of use' were correctly classified by a minimum of 70% of respondents in both groups (see Table 4.3). To achieve balance across the five indicators, only the three highest performing items tapping ‘enjoyment’ were retained. One item tapping ‘depth of use’ was retained despite not reaching the a priori threshold of 70%; as ‘depth of use’ is considered a necessary condition for engagement and one item tapping this facet was correctly classified by 65% of experts and 66% of non-experts, it was therefore considered important to retain this item. In total, ten items were retained to form the first version of the ‘DBCI Engagement Scale’ (see Table 4.4).

Table 4.3. Experts' (N = 20) and non-experts' (N = 50) classifications of the initial 18-item scale.

Item (Intended Category)	Group	Interest (%)	Attention (%)	Enjoyment (%)	Amount of use (%)	Depth of use (%)	Unclassified (%)
1. “How strongly did you experience interest?” (Interest)	Experts	**100%**	0%	0%	0%	0%	0%
	Non-experts	**84%**	4%	10%	0%	0%	2%
2. “How strongly did you experience frustration?” (Enjoyment)	Experts	5%	0%	**90%**	0%	0%	5%
	Non-experts	2%	4%	**78%**	0%	8%	8%
3. “How strongly did you experience focus?” (Attention)	Experts	0%	**95%**	0%	0%	5%	0%
	Non-experts	6%	**78%**	0%	2%	8%	6%
4. “How strongly did you experience boredom?” (Interest)	Experts	55%	10%	35%	0%	0%	0%
	Non-experts	52%	4%	38%	2%	0%	4%
5. “How strongly did you experience inattention?” (Attention)	Experts	0%	**100%**	0%	0%	0%	0%
	Non-experts	0%	**94%**	0%	4%	0%	2%
6. “How strongly did you experience absorption?” (Attention)	Experts	20%	30%	5%	0%	45%	0%
	Non-experts	18%	16%	2%	12%	46%	6%
7. “How strongly did you experience annoyance?” (Enjoyment)	Experts	5%	0%	**90%**	0%	0%	5%
	Non-experts	6%	0%	**80%**	0%	8%	6%
8. “How strongly did you experience fascination?” (Interest)	Experts	80%	0%	15%	0%	5%	0%
	Non-experts	40%	12%	32%	2%	6%	8%
9. “How strongly did you experience distraction?” (Attention)	Experts	0%	**85%**	5%	0%	10%	0%
	Non-experts	14%	**80%**	0%	2%	4%	0%

Note. Percentages in bold indicate items that were correctly classified by a minimum of 70% of respondents in both groups.

Table 4.3. *Continued.*

Item (Intended Category)	Group	Interest (%)	Attention (%)	Enjoyment (%)	Amount of use (%)	Depth of use (%)	Unclassified (%)
10. "How strongly did you experience enjoyment?" (Enjoyment)	Experts	0%	0%	**100%**	0%	0%	0%
	Non-experts	8%	0%	**84%**	2%	0%	6%
11. "How strongly did you experience intrigue?" (Interest)	Experts	**80%**	10%	0%	0%	5%	5%
	Non-experts	**74%**	6%	12%	2%	2%	4%
12. "How strongly did you experience mindfulness?" (Attention)	Experts	0%	55%	0%	0%	25%	20%
	Non-experts	8%	56%	4%	2%	22%	8%
13. "How strongly did you experience fun?" (Enjoyment)	Experts	5%	0%	**95%**	0%	0%	0%
	Non-experts	0%	4%	**86%**	4%	0%	6%
14. "How strongly did you experience pleasure?" (Enjoyment)	Experts	0%	0%	**100%**	0%	0%	0%
	Non-experts	2%	2%	**92%**	0%	2%	2%
15. "How strongly did you experience indifference?" (Interest)	Experts	60%	0%	25%	0%	0%	15%
	Non-experts	46%	4%	16%	2%	8%	22%
16. "How much time (in minutes) do you roughly think that you spent on the app?" (Amount of use)	Experts	0%	0%	0%	**85%**	5%	10%
	Non-experts	0%	2%	2%	**80%**	8%	8%
17. "Which of the app's components did you visit (e.g. diary, goal setting, game)?" (Depth of use)	Experts	0%	0%	0%	20%	65%	15%
	Non-experts	4%	4%	0%	8%	66%	18%
18. "Which component was most memorable?" (Depth of use)	Experts	15%	30%	0%	0%	15%	40%
	Non-experts	16%	18%	32%	2%	12%	20%

Note. Percentages in bold indicate items that were correctly classified by a minimum of 70% of respondents in both groups.

Table 4.4. The first version of the 'DBCI Engagement Scale'.

Please answer the following questions with regards to your most recent use of the *Drink Less* app.

How strongly did you experience the following?

1. Interest
2. Intrigue
3. Focus
4. Inattention
5. Distraction
6. Enjoyment
7. Annoyance
8. Pleasure

Measured on a 7-point scale with end-points and middle anchored: 'not at all'; 'moderately'; 'extremely'

9. How much time (in minutes) do you roughly think that you spent on the app?

Enter free text

10. Which of the app's components do you remember visiting? (You can select multiple options)

a) Calendar (Self-monitoring/feedback)
b) Create and view goals (Goal setting)
c) What has and hasn't worked (Self-monitoring/feedback)
d) Create and view action plans (Action planning)
e) Your hangover and you (Self-monitoring/feedback)
f) Review your drinking (Normative feedback)
g) Dashboard (Self-monitoring/feedback)
h) Game (Cognitive bias re-training)
i) Drink + me (Identity change)
j) Useful information (Other)
k) Other (Other)
l) Can't remember (Other)

*Indexed as a proportion of available modules (e.g. 5/7 * 100 = 71.4).*

4.3.2 Scale evaluation

A pre-registered protocol can be found on the Open Science Framework (OSF; see osf.io/qcmx4). Ethical approval was granted by UCL's Departmental Research Ethics Committee (UCLIC/1213/015).

4.3.2.1 Inclusion criteria

Participants were eligible to take part in the evaluation study if they had i) downloaded the alcohol reduction app *Drink Less* (see [95] for a detailed description of the app's content and Appendix 6 for illustrative screen shots) onto an iPhone or iPad during the study period (17th May 2017-6th March 2018); ii) not opted out from allowing their data to be used for research purposes; iii) reported being 18+ years; iv) reported residing in the UK; v) confirmed that they intended to reduce their drinking through responding "Interested in drinking less alcohol" to the question: "Why are you using *Drink Less*?"; and vi) reported an AUDIT score of 8 or more, indicating excessive alcohol consumption [77]. Eligibility was determined during the app registration process. The *Drink Less* app was selected as it includes evidence-based behaviour change techniques, it has been designed with user-input, it is freely available on the UK Apple App Store and because the researcher had access to the app's raw usage data.

4.3.2.2 Sampling

As app users are most likely to disengage after their first login session [29,30], novice users who had just downloaded the *Drink Less* app were recruited. The study was not publicly advertised. Interested participants identified the app on the Apple App Store or through word-of-mouth.

4.3.2.3 Sample size

Due to the scarcity of prior research, it was not possible to predict what parameter estimates to expect. A minimum of 200 participants was therefore planned to be recruited, as this has been recommended as a rule-of-thumb for Confirmatory Factor Analysis (CFA) [298].

4.3.2.4 Measures

In addition to the 'DBCI Engagement Scale', data were collected on: 1) gender; 2) type of work (i.e. manual, non-manual, other); and 3) location during first use of the *Drink Less* app (i.e. home, work, vehicle, public transport, restaurant/pub/café, other's home, can't remember, other).

To allow the assessment of the scale's criterion, predictive and incremental validity, app screen views were automatically recorded, stored in an online database (NodeChef) and extracted using the free python library *pandas* (https://pandas.pydata.org/) to calculate objective 'amount of use', 'depth of use' and 'number of subsequent logins'. The variable 'amount of use' was derived by calculating the time spent (in seconds) during participants' first login session. The variable 'depth of use' was operationalised as the number of app modules visited during participants' first login session, indexed as a proportion of the number of available modules (i.e. Goal Setting; Self-monitoring/Feedback; Action Planning; Normative Feedback; Cognitive Bias Re-Training; Identity Change; Other [95]). A new login was defined as a new screen view after 30 minutes of inactivity [302]. Participants were also asked to respond to the two 'best bets' for a short measure of engagement (described above).

To allow the assessment of the scale's divergent validity, participants were asked to respond to two items tapping the state of 'flow', as this was conceptualised as a qualitatively distinct state [43]. Although engagement with DBCIs is expected to share some experiential qualities with the state of flow (i.e. 'attention', 'interest'), it was expected that users will not necessarily experience 'balance between challenge and skill' or 'loss of time and self-consciousness' when engaging with a DBCI. Therefore, assessing whether users can experience engagement without necessarily experiencing the state of flow was considered a useful test of the scale's divergent validity. Two items from the 'Flow State Scale' [47], measured on 5-point Likert scales, that had previously been found to load most strongly onto the general 'flow' factor were selected (i.e. "When using *Drink Less*, the way time passed seemed to be different from normal"; "When using *Drink Less*, I was not worried about what others may have been thinking of me"). Although the original 'Flow State Scale' is made up of 36 items, only two of its most strongly loading items were included to minimise measurement burden.

4.3.2.5 Procedure

Eligible participants were prompted to fill out the 'DBCI Engagement Scale' immediately after their first login session. Use of the smartphone's home button to exit *Drink Less* triggered a local push notification with a link to the scale. Participants were asked to read the information sheet and provide informed consent prior to completing the scale. The push notification contained the following message: "Help science by responding to a brief survey." Due to slow recruitment (i.e. ~3 responses/week), the message was changed on the 9th August 2017 to: "Take a brief survey and enter a prize draw to win one of thirty

£10 Amazon vouchers”. This incentive was chosen as the literature suggests that participants in online surveys respond at least as well to prize draws as other incentives [303]. This resulted in an average response rate of 5.5 responses/week, although it should be noted that this time period included the New Year period, in which there was an isolated spike of responses.

4.3.3 Data analysis

All analyses were conducted using SPSS version 20.0 [304]. The assumptions for parametric tests were assessed (i.e. normality of the distribution of residuals). When violated, normalisation was used (i.e. *z*-normalisation for positively skewed data). Descriptive statistics (e.g. mean, range, variance) were calculated for each of the scale items and the additional variables of interest to determine suitability for factor analysis.

4.3.3.1 Construct validity

It was hypothesised that a five-factor solution (i.e. ‘amount of use’, ‘depth of use’, ‘interest’, ‘attention’, ‘enjoyment’) would provide the best fit of the observed data. Pre-planned analyses registered on the OSF therefore included the use of CFA. However, due to potential range restriction in key outcome variables resulting from self-selection during the recruitment process (i.e. only a small number of eligible users completing the scale), Exploratory Factor Analysis (EFA) using principal axis factoring and oblique rotation was deemed more suitable. Inspection of Cattell’s scree plots and the Kaiser criterion (i.e. eigenvalues > 1) were used to determine the number of factors to retain [305]. Pre-planned analyses also included a comparison of the fit of the CFA solution using the self-reported data as input with a CFA solution using a combination of self-reported data (i.e. the experiential indicators) and automatically recorded usage data (i.e. the behavioural indicators). However, an additional EFA was deemed more suitable.

4.3.3.2 Internal consistency reliability

Internal consistency reliability was assessed by calculating Cronbach's α. A large coefficient (i.e. α = .70 or above) was interpreted as evidence of strong item covariance [298].

4.3.3.3 Criterion validity

Criterion validity was assessed by calculating Pearson's correlation coefficient for the relationship between participants' automatically recorded screen views from their first login session (i.e. objective 'amount of use' and 'depth of use') with the self-reported scale items (i.e. subjective 'amount of use' and 'depth of use'), and with participants' total scale scores.

4.3.3.4 Predictive validity

Pre-planned analyses registered on the OSF included a regression analysis in which the outcome variable 'subsequent login' (i.e. whether or not participants ever logged in again) would be regressed onto total scale scores. As all but 3.4% (7/203) of participants returned to the app after their first login session, this variable would have failed to discriminate between participants. Instead, an unplanned analysis was conducted, in which the variable 'number of subsequent logins', operationalised as the total number of logins in the 14 days after app registration, was regressed onto total scale scores. A cut-off at 14 days post-registration was deemed appropriate as DBCI access tends to be most prevalent during this time window [306].

4.3.3.5 Incremental validity

Incremental validity was assessed through examining the additional variance accounted for in ‘number of subsequent logins’ after adding the self-reported experiential indicators (but not the self-reported behavioural indicators) to a model including only the objectively recorded behavioural indicators of engagement.

4.3.3.6 Divergent validity

Divergent validity was assessed by calculating Pearson’s correlation coefficient for the relationship between each of the two items tapping the state of ‘flow’ and the overall measure of engagement.

4.3.3.7 Unplanned sensitivity analyses

As only a small proportion of eligible participants completed the scale, an unplanned sensitivity analysis was required to examine whether there was potential range restriction in the scale items and key outcome variables. A Mann-Whitney *U* test was conducted to assess whether the median number of subsequent logins differed significantly between those who did and did not complete the scale. An additional unplanned sensitivity analysis was conducted to assess if participants’ AUDIT scores were significantly associated with total scale scores or the number of subsequent logins.

4.4 Results

4.4.1 Participants

During the study period (294 days; 17th May 2017-6th March 2018), a total of 8,336 users downloaded the *Drink Less* app, of which 5,460 (65.5%) were eligible to complete the scale. Of these, 311 (5.7%) users initiated the scale (i.e. opened the link), with 203 (3.7%) users completing it (see Figure 4.1). Participant demographic and drinking characteristics are reported in Table 4.5.

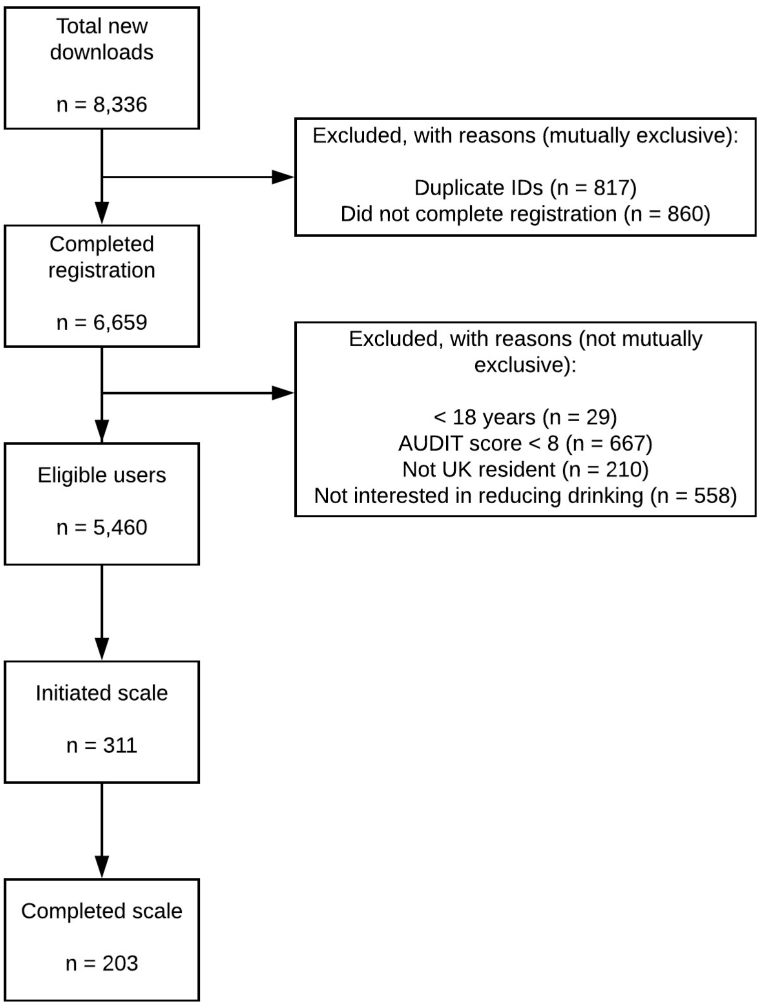


Figure 4.1. Participant flow chart.

Table 4.5. Participants' demographic and drinking characteristics.

Demographic characteristics	**Completed scale** (***N*** **= 203)**	**Initiated (but not completed) scale** (***N*** **= 108)**	p^a
Female, % (*N*)	64% (129)	53% (57)	.07
Type of work, % (*N*)			.85
Non-Manual, % (*N*)	75% (152)	73% (79)	
Manual, % (*N*)	11% (22)	11% (11)	
Other, % (*N*)	14% (29)	17% (18)	
Age in years, mean (*SD*)	41.8 (*10.7*)	42.4 (*9.5*)	.66
Drinking characteristics			
AUDIT, mean (*SD*)	17.6 (*6.1*)	18.3 (*6.8*)	.31

Note. [a] Differences between groups were assessed using chi-square tests or *t*-tests, as appropriate.

4.4.2 Descriptive statistics for scale items

Descriptive statistics for the scale items are reported in Table 4.6. The majority of participants completed the scale at home (83%) or at work (7.9%). To account for observed skewness, *z*-score transformation was applied to the 10 scale items, the two items used for the criterion validity analyses and the three items used for the predictive and incremental validity analyses. Inter-item correlations of the normalised items are reported in Table 4.7.

Table 4.6. Descriptive statistics for the scale items (*N* = 203)

	Range	Mean (*SD*)	Variance	Skewness	Kurtosis
Scale Items					
1. “How strongly did you experience interest?”	1-7	5.43 (*1.19*)	1.41	-0.61	0.34
2. “How strongly did you experience intrigue?”	1-7	5.05 (*1.57*)	2.48	-0.70	-0.27
3. “How strongly did you experience focus?”	2-7	5.06 (*1.24*)	1.54	-0.52	-0.07
4. “How strongly did you experience inattention?” (R)	1-7	5.32 (*1.47*)	2.17	-0.92	0.22
5. “How strongly did you experience distraction?” (R)	1-7	5.30 (*1.65*)	2.72	-0.91	-0.13
6. “How strongly did you experience enjoyment?”	1-7	4.30 (*1.40*)	1.95	-0.31	-0.37
7. “How strongly did you experience pleasure?”	1-7	3.63 (*1.56*)	2.44	0.07	-0.85
8. “How strongly did you experience annoyance?” (R)	1-7	5.77 (*1.40*)	1.97	-1.27	1.10
9. “Which of the app’s components did you visit?”	14.29-100.00	53.34 (*22.99*)	528.97	0.20	-0.73
10. “How much time do you roughly think that you spent on the app?” (seconds)	120-3600	561.87 (*379.07*)	143,697.47	3.62	22.65
Items used to test Criterion Validity					
11. Objective depth of use	14.29-100.0	77.62 (*16.69*)	278.68	-0.66	0.40
12. Objective amount of use (seconds)	0-3303	802.57 (*646.03*)	417,354.87	1.96	3.98
Items used to test Predictive/Incremental Validity					
13. Number of subsequent logins	0-67	15.40 (*12.35*)	152.51	1.39	2.66
14. “How engaging was the app?”	1-7	5.15 (*1.16*)	1.34	-0.83	1.39
15. “How much did you like the app?”	2-7	5.33 (*1.11*)	1.23	-0.50	-0.14
Items used to test Divergent Validity					
16. “When using Drink Less, the way time passed seemed different from normal.”	1-5	2.87 (*0.73*)	0.53	-0.56	0.93
17. “When using Drink Less, I was not worried about what others may have been thinking about me.”	1-5	2.78 (*1.21*)	1.47	0.10	-1.05

Note. The symbol (R) indicates that values have been reverse scored prior to the calculation of descriptive statistics.

Table 4.7. Inter-item correlation matrix (*N* = 203).

Scale Items	**1. Interest**	**2. Intrigue**	**3. Focus**	**4. Inattention (R)**	**5. Distraction (R)**	**6. Enjoyment**	**7. Pleasure**	**8. Annoyance (R)**	**9. Which of app’ s components**	**10. How much time spent**
1. Interest	1									
2. Intrigue	.56***	1								
3. Focus	.73***	.53***	1							
4. Inattention (R)	.11	-.00	.14	1						
5. Distraction (R)	.10	-.02	.14*	.60***	1					
6. Enjoyment	.43***	.57***	.40***	-.08	-.15*	1				
7. Pleasure	.29***	.41***	.31***	-.25***	-.30***	.62***	1			
8. Annoyance (R)	.24***	.14*	.27***	.43***	.44***	.09	.02	1		
9. Which of app’s components	.17*	.27***	.13	.03	.04	.20***	.14*	.11	1	
10. How much time spent	.19***	.16*	.13	-.23***	-.17*	.16*	.19***	-.06	.28***	1

Note. The symbol (R) indicates that values have been reverse scored prior to analysis; * $p < .05$; ** $p < .01$; *** $p < .001$.

4.4.3 Scale validity

4.4.3.1 Construct validity

The Keiser-Meier Olkin (KMO) Test of Sampling Adequacy (KMO = 0.76) and Bartlett's Test of Sphericity (p < .001) indicated that the data were suited for factor analysis [307]. Three different EFA solutions were subsequently tested to arrive at a best fitting solution.

4.4.3.1.1 Solution 1

An EFA using principal axis factoring estimation with oblique rotation indicated that a two-factor solution, accounting for 54.1% of the variance, provided the best fit of the observed data (see Table 4.8). However, the second factor was comprised only of the negatively worded items (i.e. items 4, 5 and 8), making little theoretical sense. On the basis of the conceptual parsimony of a one-factor solution, the second factor and the negatively worded items were discarded.

4.4.3.1.2 Solution 2

A subsequent EFA with oblique rotation indicated that a one-factor solution accounted for 44.5% of the variance (see Table 4.8).

4.4.3.1.3 Solution 3

An EFA with oblique rotation using a combination of self-reported data (i.e. items 1, 2, 3, 6, and 7) and automatically recorded data (i.e. items 11 and 12) suggested that a two-factor solution provided an adequate fit, which accounted for 63.5% of the variance. The experiential indicators loaded clearly onto factor 1, and the behavioural indicators loaded clearly onto factor 2 (see Table 4.8).

Table 4.8. Factor loadings of the ‘DBCI Engagement Scale’ in EFAs.

Scale Items	**Solution 1***		**Solution 2****	**Solution 3*****	
	Factor 1	**Factor 2**	**Factor 1**	**Factor 1**	**Factor 2**
1. Interest	**0.51**	0.14	**0.53**	**0.74**	0.05
2. Intrigue	**0.65**	0.02	**0.67**	**0.76**	0.07
3. Focus	**0.49**	0.20	**0.50**	**0.72**	0.01
4. Inattention (R)	-0.15	**0.79**	N/A	N/A	N/A
5. Distraction (R)	-0.21	**0.78**	N/A	N/A	N/A
6. Enjoyment	**0.86**	-0.09	**0.85**	**0.71**	-0.03
7. Pleasure	**0.60**	-0.32	**0.70**	**0.56**	-0.09
8. Annoyance (R)	0.11	**0.56**	N/A	N/A	N/A
9. Which of app’s components	**0.26**	0.03	**0.28**	N/A	N/A
10. How much time spent	**0.25**	-0.23	**0.26**	N/A	N/A
11. Objective depth of use	N/A	N/A	N/A	0.10	**0.73**
12. Objective amount of use	N/A	N/A	N/A	-0.09	**0.64**

Note. The symbol (R) indicates that values have been reverse scored prior to analysis. * EFA with oblique rotation, including items 1-10; ** EFA with oblique rotation, including items 1, 2, 3, 6, 7, 9 and 10; *** EFA with oblique rotation, including items 1, 2, 3, 6, 7, 11 and 12

Solution 2 was selected for further analyses, as it contained only the self-reported items. A total scale score was calculated for each participant, with equal weight given to each of the retained self-reported items (i.e. items 1, 2, 3, 6, 7, 9 and 10).

4.4.3.2 Internal consistency reliability

Internal consistency estimates for the 7-item scale yielded a coefficient α of .77, indicating adequate internal consistency reliability [301].

4.4.3.3 Criterion validity

Total scale scores were significantly correlated with objectively recorded 'depth of use', $r(201) = 0.23$, $p < .01$, but not with objectively recorded 'amount of use', $r(201) = -0.02$, $p = .82$. Self-reported 'depth of use' was significantly correlated with objectively recorded 'depth of use', $r(201) = 0.44$, $p < .001$. Self-reported 'amount of use' was significantly correlated with objectively recorded 'amount of use', $r(201) = 0.15$, $p < .05$.

4.4.3.4 Predictive validity

The overall measure did not predict the number of subsequent logins (B = .02, 95% CI = -.01, .05, $p = .14$). Asking users about how engaging they thought the app was or how much they liked the app did not predict the number of subsequent logins (see Table 4.9). A post-hoc power analysis indicated that a total of 203 participants provided 44% power (two-tailed $\alpha = .05$) to detect a regression coefficient of .02 for the association between total scale scores and the number of subsequent logins [39] (although it must be noted that post-hoc power analyses should be interpreted with caution [308]).

Table 4.9. Univariate and multivariate linear regression models predicting the number of subsequent logins.

	B (95% CI)	*p*-value
Predictive Validity		
Total scale scores	.02 (-.01, .05)	.14
How engaging was the app?	.07 (-.07, .21)	.30
How much did you like the app?	.09 (-.05, .22)	.20
Incremental Validity		
Model 1		
Objective amount of use	.07 (-.09, .22)	.40
Objective depth of use	.03 (-.13, .18)	.75
Model 2		
Objective amount of use	.09 (-.07, .25)	.27
Objective depth of use	-.01 (-.17, .15)	.89
Interest	.25 (.03, .46)	.02*
Focus	-.10 (-.30, .11)	.35
Enjoyment	.02 (-.18, .22)	.86
Intrigue	.04 (-.15, .22)	.71
Pleasure	-.02 (-.20, .15)	.79

* $p < .05$

4.4.3.5 Incremental validity

Results from the regression analyses are reported in Table 4.9. A model including the automatically recorded indicators of engagement (i.e. items 11 and 12) accounted for 0.7% of variance in the number of subsequent logins (Model 1). Neither objective ‘amount of use’ nor objective ‘depth of use’ were significant predictors of the number of subsequent logins. A model including the automatically recorded indicators in addition to the experiential indicators of engagement (i.e. items 1, 2, 3, 6 and 7) accounted for 4.9% of variance in the number of subsequent logins (Model 2). Interest was the only significant predictor of the number of subsequent logins.

4.4.3.6 <u>Divergent validity</u>

Total scale scores were significantly correlated with the first ("When using *Drink Less*, the way time passed seemed different from normal") but not the second ("When using *Drink Less*, I was not worried about what others may have been thinking about me") item tapping flow ($r(201) = 0.14$, $p = .04$ and $r(201) = -0.07$, $p = .33$, respectively). The two items tapping flow were not significantly correlated with each other in this sample ($r(201) = -0.02$, $p = .82$).

4.4.3.7 <u>Unplanned sensitivity analyses</u>

The unplanned sensitivity analysis indicated that those who completed the scale had a significantly greater median number of subsequent logins (median = 13.0, interquartile range (IQR) = 6.0-21.0) than eligible users who did not complete the scale (median = 6.0, IQR = 1.0-16.0), $U = 361{,}135.5$, $p < .001$. The second sensitivity analysis showed that participants' AUDIT scores were neither significantly correlated with total scale scores ($r(201) = .10$, p = .14), nor with the number of subsequent logins ($r(201) = .004$, p = .95).

4.5 Discussion

4.5.1 Summary of key findings

This study described the systematic development of a new self-report measure of engagement with DBCIs and its validation in a real-world setting with an alcohol reduction app. As fewer than 5% of eligible users completed the scale, the first observation is that it was not established that it is feasible to measure engagement through self-report in a real-world setting. Secondly, results from a series of EFAs indicate that the 7-item ‘DBCI Engagement Scale’ is unifactorial and internally reliable. Thirdly, total scale scores were significantly but weakly correlated with objective ‘depth of use’ but not significantly correlated with objective ‘amount of use’, thus questioning the scale’s criterion validity. Fourthly, total scale scores did not predict the number of subsequent logins in the next 14 days. Finally, total scale scores were significantly correlated with one of the two items from the Flow State Scale, thus questioning the scale’s divergent validity.

4.5.2 Limitations

These results should be interpreted in the light of a number of important methodological and theoretical limitations. Through comparing the number of subsequent logins between the analytic sample and the sample of total eligible users, it is evident that the analytic sample was biased towards highly engaged users. It is likely that this restricted the range in both scale items and key outcome variables, thus limiting the ability of the present study to evaluate the scale’s validity. The inclusion criteria (i.e. expressing a desire to reduce drinking, being willing to use an app, being willing to share data with the researchers) may also have contributed to the apparent self-selection bias. However, these inclusion criteria mirror those in randomised controlled trials of health apps [31, 40]. It is notoriously difficult to study engagement in real-world settings, as highly engaged individuals are more likely to take part in such research (i.e. users who login more frequently have a greater likelihood of responding to follow-up surveys) [153]. An important avenue for future research is therefore to evaluate the scale's validity in a more controlled setting, with a view to recruiting participants with a broader range of engagement levels (e.g. students or participants taking part in research for credit or financial rewards).

The observation that the negatively worded items (e.g. 'inattention', 'distraction') were found to load onto a second factor in the initial EFA (which resulted in the removal of these items) suggests that participants may have found it difficult to respond to the negatively worded items. Despite having assessed the items' content validity through an initial content adequacy task, it is possible that 'inattention' is not seen as the polar opposite of 'attention' in everyday language. Future work using cognitive interviewing techniques is therefore required to refine the scale items, ensuring that the retained items are easy to respond to [309]. Moreover, the observation that the two items assessing the state of flow were not significantly correlated in this sample also highlights the importance of using well-validated scales when benchmarking a new scale, where available.

The lack of an association between initial experiential and behavioural engagement and future engagement can be interpreted in multiple ways. First, the study was not powered to detect a weak relationship between initial and future engagement. Secondly, it is plausible that other factors, such as motivation to change the target behaviour or perceived personal relevance, are in fact more strongly predictive of future engagement than initial experiential and behavioural engagement. Indeed, systematic reviews of DBCIs indicate that aggregate measures of engagement (e.g. total number of logins over a period of time) are influenced by attributes of the DBCI itself (e.g. tailoring, aesthetics), characteristics of the users (e.g. motivation to change) and the context in which the DBCI is used (e.g. social cues) [281,288]. It was therefore decided that this should be examined further in Chapter 5.

4.5.3 Conclusion

Behavioural and experiential indicators of engagement may resolve to a single dimension. Low response rates to engagement surveys embedded in DBCIs may make their use impracticable in real-world settings. The lack of an association between total scale scores and the number of subsequent logins suggests that other factors, such as motivation to change, may play a more important role in the prediction of future engagement than initial behavioural and experiential engagement.

4.5.4 Citation for the peer-reviewed article for this study

Perski, O., Blandford, A., Garnett, C., Crane, D., West, R., & Michie, S. (under review). A self-report measure of engagement with digital behaviour change interventions (DBCIs): Development and psychometric evaluation of the ‘DBCI Engagement Scale’. *Translational Behavioral Medicine.*

4.5.5 Next steps

The next step of the thesis was to evaluate the ‘DBCI Engagement Scale’ in a different population with a potentially broader range of engagement levels whilst also taking account of users’ motivation to change (reported in Chapter 5).

5 CHAPTER 5 – On the dimensional structure of engagement with digital behaviour change interventions (DBCIs): Psychometric evaluation of the ‘DBCI Engagement Scale’ in a new population (Study 4)

5.1 Abstract

Background: The ‘DBCI Engagement Scale’ was designed to capture the behavioural and experiential dimensions of engagement with DBCIs. Results from an initial evaluation study suggested that these indicators of engagement may resolve to a single dimension; however, low response rates to the survey and range restriction in both scale items and key outcome variables limited efforts to evaluate the scale’s psychometric properties.

Purpose: The present study aimed to evaluate the ‘DBCI Engagement Scale’ in a population with a broad range of engagement levels.

Methods: The study sample consisted of UK-based adults who drink excessively and were willing to download the *Drink Less* app and complete the scale immediately after their first login in exchange for a financial reward, recruited via ‘Prolific’, an online research platform. Five types of validity (i.e. construct, criterion, predictive, incremental, divergent) were examined using factor analysis, correlational analyses, and regression analyses. Cronbach’s α was calculated to assess internal reliability.

Results: Of 266 eligible participants, 147 (55%) completed the scale. Six items were retained. A two-factor solution, with the experiential indicators loading onto factor 1 (‘Experiential Engagement’) and the behavioural indicators loading onto

factor 2 ('Behavioural Engagement'), provided the best fit. The scale did not show good internal consistency (α = .67), nor divergent and criterion validity. Total scale scores predicted the variable 'subsequent login' in both unadjusted and adjusted models controlling for motivation to reduce alcohol consumption (OR_{adj} = 1.14, 95% CI = 1.03-1.27, p = .01).

Conclusion: Experiential and behavioural indicators of engagement may constitute two separate dimensions. The overall measure predicted future behavioural engagement. This remained significant when adjusting for motivation to reduce alcohol. Due to not achieving the desired sample size, these findings merit replication in a larger sample.

5.2 Introduction

As outlined in Chapter 4, results from the initial evaluation of the 'DBCI Engagement Scale' suggested that it had provided a suboptimal test of the scale's psychometric properties. It is notoriously difficult to study engagement in real-world settings, as highly engaged individuals are more likely to take part in such research (e.g. users who login to DBCIs more frequently have a greater likelihood of responding to follow-up surveys) [99]. It was therefore considered important to evaluate the scale's psychometric properties in a different setting, with a view to recruiting participants with a potentially broader range of engagement levels. As motivation to reduce alcohol consumption tends to be positively associated with behavioural engagement [164,310], it was considered important to adjust for this variable in predictive validity analyses. The aim of the present study was to evaluate the 'DBCI Engagement Scale' in a sample of UK-based, adult, excessive drinkers recruited via ‘Prolific’ [311], an online, web-based platform for recruiting and paying participants to complete tasks.

5.3 Methods

5.3.1 Study design

A pre-registered study protocol can be found on the OSF (osf.io/qcmx4). Ethical approval was granted by UCL’s Computer Science Departmental Research Ethics Chair (Project ID: UCLIC/1617/004/Staff Blandford HFDH).

5.3.1.1 Inclusion criteria

Participants were eligible to take part if they i) were aged 18+ years; ii) reported an AUDIT score of ≥ 8, indicating excessive alcohol consumption [77]; iii) resided in the UK; iv) owned an iPhone capable of running iOS 8.0 software (i.e. an iPhone 4S or later models); and v) were willing to download and explore an alcohol reduction app (see Appendix 7).

5.3.1.2 Sampling

Participants were recruited via ‘Prolific’ (www.prolific.ac.uk) [311]. As app users are most likely to disengage with health apps after their first login session [29,30], novice users were recruited.

5.3.1.3 Sample size

As is commonly specified in the psychometric literature, a 25:1 participant-to-item ratio (i.e. a total of 250 participants) was considered desirable [305].

5.3.1.4 Measures

During the screening phase, data were collected on: i) age; ii) gender; iii) type of work (i.e. manual, non-manual, other); iv) patterns of alcohol consumption, measured by the AUDIT; v) motivation to reduce alcohol consumption, measured by MTSS [262,265,312]; and vi) willingness to download and explore an alcohol reduction app (yes vs. no).

After having downloaded and explored the *Drink Less* app, data were collected on: i) location during first use of the app (i.e. home, work, vehicle, public transport, restaurant/pub/café, other's home, can't remember, other); ii) the 10-item 'DBCI Engagement Scale'; iii) two items from the 'Flow State Scale' that have previously been found to load most strongly onto the general 'flow' factor (i.e. "When using *Drink Less*, the way time passed seemed to be different from normal"; "When using *Drink Less,* I was not worried about what others may have been thinking of me"); and iv) two items that represent the author's best bets for a short measure of engagement (i.e. "How much did you like the app?"; "How engaging was the app?").

App screen views were recorded automatically during participants' first login session and continuously over the next few days (to be able to derive the variable 'subsequent login', described below). App screen views were stored in an online database (NodeChef) and extracted using the free python library *pandas* (https://pandas.pydata.org/) to derive objective 'amount of use', 'depth of use' and the variable 'subsequent login'. The variable 'amount of use' was derived by calculating the time spent (in seconds) during participants' first login session. The variable 'depth of use' was derived by calculating the number of app components visited during participants' first login session, indexed as a proportion of the number of available components (i.e. Goal Setting; Self-monitoring/Feedback; Action Planning; Normative Feedback; Cognitive Bias Re-Training; Identity Change; Other [95]). The variable 'subsequent login' was derived by assessing whether participants had accessed the app again after their first login session (no vs. yes). A new login session was defined as a new screen view following at least 30 minutes of inactivity [302].

5.3.1.5 __Procedure__

Interested participants, identified via Prolific's online platform, were asked to complete the screening questionnaire to assess study eligibility (see Appendix 7). The screening questionnaire was hosted by Qualtrics survey software [266]. Participants were paid £0.5 for completing the screening questionnaire.

Eligible participants were sent an e-mail invitation via Prolific to download the *Drink Less* app from the Apple App Store and to explore it as they would explore any new app (see Appendix 8). Participants were told that the researcher would monitor their usage of the app to assess what content they were interested in. For technical reasons, participants were told that they had to select the option 'interested in drinking less alcohol' when asked about why they were using the *Drink Less* app and to enable push notifications, as they would otherwise not be sent the link to the study survey. When clicking on their phone's home button after having finished exploring the app, participants received a push notification with a link to the survey. After completing this, they were asked to enter their Prolific ID number, which enabled the researcher to provide payment to participants and to match participants' survey responses to their screen view records. All participants who initiated but did not complete the task (as indicated by their response status on the Prolific platform, which was either labelled ‘Timed out’ or ‘Returned submission’), were sent one reminder message by the researcher. Participants were paid £1.25 for completing the task.

5.3.2 Data analysis

All analyses were conducted using SPSS version 20.0 [304]. The assumptions for parametric tests were assessed (i.e. normality of the distribution of residuals) and when violated, normalisation was used (i.e. *z*-normalisation for positively skewed data). Descriptive statistics (e.g. mean, range, variance) were calculated for each of the scale items and the additional variables of interest to determine suitability for factor analysis.

5.3.2.1 <u>Construct validity</u>

It was hypothesised that a five-factor solution (i.e. ‘amount of use’, ‘depth of use’, ‘interest’, ‘attention’, ‘enjoyment’) would provide the best fit of the observed data. A series of Exploratory Factor Analyses (EFAs) using principal axis factoring estimation and oblique rotation were conducted. The inspection of Cattell’s scree plots and the Kaiser criterion (i.e. eigenvalues > 1) were used to determine the number of factors to retain [305].

5.3.2.2 Internal consistency reliability

Internal consistency reliability was assessed through calculating Cronbach's α. A large coefficient (i.e. α = .70 or above) was interpreted as evidence of strong item covariance.

5.3.2.3 Criterion validity

Criterion validity was assessed by calculating Pearson's correlation coefficient for the relationship between participants' automatically recorded screen views from their first login session (i.e. objective 'amount of use' and 'depth of use') with the self-reported scale items (i.e. subjective 'amount of use' and 'depth of use'), and with participants' total scale scores.

5.3.2.4 Predictive validity

The analysis of the scale's predictive validity proceeded in several steps. The variable 'subsequent login' was first regressed onto participants' total scale scores in an unadjusted logistic regression analysis. The variable 'subsequent login' was then regressed onto each of the two 'best bets' for a short measure of engagement (i.e. "How engaging was the app?"; "How much did you like the app?"). The univariate association between motivation to reduce alcohol consumption and the variable 'subsequent login' was subsequently assessed. As motivation to reduce alcohol consumption was found to be significantly associated with the variable 'subsequent login', a series of multivariate logistic regression analyses were conducted, repeating the above univariate analyses, adjusting for motivation to reduce alcohol.

5.3.2.5 Incremental validity

Incremental validity was assessed by examining the additional variance accounted for in the variable ‘subsequent login’ after adding the self-reported experiential indicators to a model including only the automatically recorded behavioural indicators of engagement.

5.3.2.6 Divergent validity

Divergent validity was assessed through calculating Pearson’s correlation coefficient for the relationship between each of the two items tapping the state of ‘flow’ and the overall measure of engagement.

5.4 Results

5.4.1 Participants

During the study period (31 days; 23rd July 2018-22nd August 2018), a total of 401 participants completed the online screening survey, of which 266 were eligible to download the *Drink Less* app and respond to the survey. Of these, 147 (55%) completed the task (see Figure 5.1). The desired target of 250 participants was hence not achieved. Four participants who had initiated but not completed the task responded to the messages indicating that they had experienced technical issues due to i) already having installed the *Drink Less* app prior to the study (n = 1), ii) failing to install the app despite trying multiple times (n = 1), iii) receiving an error message during the onboarding stage which hindered the participant from proceeding any further (n = 1), and iv) not receiving the push notification containing the survey on their iPhone X, as it did not have a home button (although the developers confirmed this should not have prevented the notification to be delivered) (n = 1). Participant demographic and drinking characteristics are reported in Table 5.1.

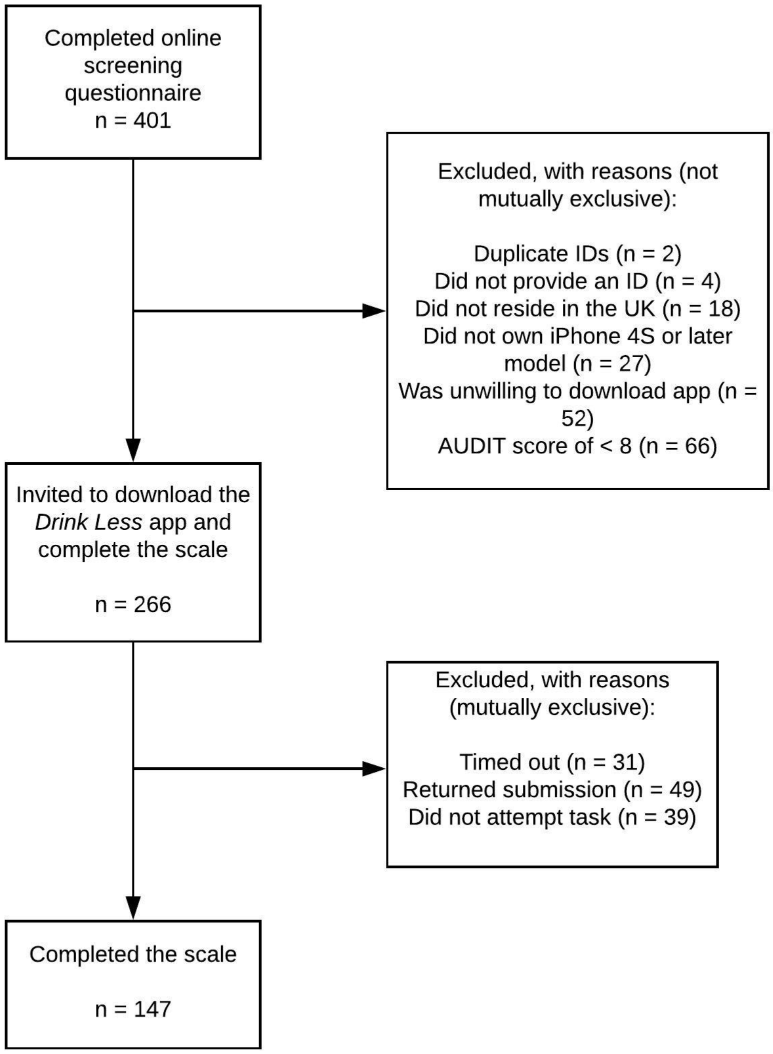


Figure 5.1. Participant flow chart.

Table 5.1. Participant demographic and drinking characteristics.

Demographic characteristics	**Completed scale (*N* = 147)**	**Eligible but did not complete scale (*N* = 119)**	***p*[a]**
Female, % (*N*)	66% (97)	60% (71)	.29
Type of work, % (*N*)			.57
Manual, % (*N*)	13% (19)	13% (16)	
Non-Manual, % (*N*)	61% (89)	66% (78)	
Other, % (*N*)	27% (39)	21% (25)	
Age in years, mean (*SD*)	34.4 (*10.4*)	36.6 (*11.8*)	.11
Drinking characteristics			
MTSS*			.08
1) I don’t want to cut down on drinking alcohol	10% (14)	22% (26)	
2) I think I should cut down on drinking alcohol but I don’t really want to	29% (43)	21% (25)	
3) I want to cut down on drinking alcohol but I haven’t thought about when	13% (19)	14% (17)	
4) I really want to cut down on drinking alcohol but I don’t know when I will	12% (17)	9% (11)	
5) I want to cut down on drinking and hope to soon	16% (23)	14% (17)	
6) I really want to cut down on drinking alcohol and intend to in the next 3 months	7% (11)	3% (4)	
7) I really want to cut down on drinking alcohol and intend to in the next month	14% (20)	16% (19)	
AUDIT** score, mean (*SD*)	15.4 (*5.1*)	14.2 (*5.7*)	.07

[a] Differences between groups were assessed using chi-square tests or *t*-tests, as appropriate; * MTSS = Motivation To Stop Scale; ** AUDIT = Alcohol Use Disorders Identification Test

5.4.2 Descriptive statistics

Descriptive statistics for the scale items are reported in Table 5.2. The majority of participants completed the scale at home (80.3%) or at work (12.9%).

Table 5.2. Descriptive statistics for the scale items (*N* = 147).

	Range	Mean (*SD*)	Variance	Skewness	Kurtosis
Scale Items					
1. “How strongly did you experience interest?”	2-7	5.30 (*1.09*)	1.18	-0.30	0.06
2. “How strongly did you experience intrigue?”	1-7	5.39 (*1.27*)	1.61	-0.85	0.50
3. “How strongly did you experience focus?”	2-7	5.31 (*1.18*)	1.40	-0.56	0.14
4. “How strongly did you experience inattention?” (R)	1-7	5.61 (*1.33*)	1.76	-1.24	1.47
5. “How strongly did you experience distraction?” (R)	1-7	5.47 (*1.45*)	2.10	-1.12	0.86
6. “How strongly did you experience enjoyment?”	1-7	4.46 (*1.44*)	2.07	-0.10	-0.48
7. “How strongly did you experience pleasure?”	1-7	3.56 (*1.64*)	2.67	0.36	-0.70
8. “How strongly did you experience annoyance?” (R)	1-7	5.59 (*1.39*)	1.93	-1.09	1.08
9. “Which of the app’s components did you visit?”	14.29-100.00	58.70 (*22.00*)	484.01	-0.12	-0.67
10. “How much time do you roughly think that you spent on the app?” (seconds)	120-1200	520.82 (*237.21*)	56,267.82	0.93	0.96
Items used to test Criterion Validity					
11. Objective depth of use	28.57-100.00	66.66 (*20.50*)	420.28	-0.23	-0.85
12. Objective amount of use (seconds)	95-3571	409.45 (*360.71*)	130,116.72	5.13	40.34
Items used to test Divergent Validity					
13 “When using Drink Less, the way time passed seemed different from normal.”	1-5	2.76 (*0.79*)	0.62	0.11	0.10
14. “When using Drink Less, I was not worried about what others may have been thinking about me.”	1-5	3.34 (*1.16*)	1.35	-0.24	-1.11
Items used to test Predictive/Incremental Validity					
15. “How much did you like the app?”	1-7	5.14 (*1.29*)	1.66	-0.80	0.82
16. “How engaging was the app?”	1-7	5.20 (*1.17*)	1.37	-0.65	0.66
	% (*N*)				
17. Subsequent login (yes vs. no)	46% (*67*)				

Note. The symbol (R) indicates that values have been reverse scored prior to the calculation of descriptive statistics.

To account for observed skewness, *z*-score transformation was subsequently applied to the 10 scale items and the two items used to test the scale's criterion validity. Inter-item correlations of the normalised scale items are reported in Table 5.3.

Table 5.3. Inter-item correlation matrix (*N* = 147).

Scale Items	**1. Interest**	**2. Intrigue**	**3. Focus**	**4. Inattention (R)**	**5. Distraction (R)**	**6. Enjoyment**	**7. Pleasure**	**8. Annoyance (R)**	**9. Which of app’ s components**	**10. How much time spent**
1. Interest	1									
2. Intrigue	.44***	1								
3. Focus	.65***	.46***	1							
4. Inattention (R)	.18*	.10	.31***	1						
5. Distraction (R)	.18*	.12	.28***	.43***	1					
6. Enjoyment	.48***	.31***	.44***	-.05	-.15	1				
7. Pleasure	.19*	.09	.15	-.19*	-.24***	.54***	1			
8. Annoyance (R)	.28***	.15	.37***	.27***	.34***	.29***	.12	1		
9. Which of app’s components	.18*	.00	.06	.13	-.03	.19*	.19*	.13	1	
10. How much time spent	.10	.10	-.03	.08	.11	.15	.33***	.09	.29***	1

Note. The symbol (R) indicates that values have been reverse scored prior to analysis; * *p* < .05; ** *p* < .01; *** *p* < .001.

5.4.3 Scale validity

5.4.3.1 <u>Construct validity</u>

The Keiser-Meier Olkin (KMO) Test of Sampling Adequacy (KMO = 0.70) and Bartlett's Test of Sphericity ($p < .001$) indicated that data were suited for factor analysis. Three different EFA solutions were subsequently tested to arrive at a best fitting solution.

5.4.3.1.1 Solution 1

An EFA with oblique rotation was conducted, and the eigenvalues indicated that a three-factor solution, accounting for 61.2% of the variance, was most appropriate (see Table 5.4). The loadings indicated that the second factor comprised two of the negatively worded items (4 and 5). The third factor comprised the two behavioural indicators of engagement (i.e. items 9 and 10) and one of the experiential indicators (i.e. item 7), which made little theoretical sense. The loading of item 8 (also a negatively worded item) onto factor 1 was modest. Therefore, the negatively worded items (i.e. items 4, 5 and 8) and item 7 were discarded.

5.4.3.1.2 Solution 2

A subsequent EFA with oblique rotation indicated that a two-factor solution accounted for 62.4% of the variance (see Table 5.4). The experiential indicators loaded clearly onto factor 1, and the behavioural indicators loaded clearly onto factor 2, with no cross-loadings (i.e. items that load at 0.32 or higher on two or more factors) [305]. The two latent factors were labelled ‘Experiential Engagement’ and ‘Behavioural Engagement’, respectively.

5.4.3.1.3 Solution 3

An EFA with oblique rotation using a combination of self-reported data (i.e. items 1, 2, 3 and 6) and automatically recorded data (i.e. items 11 and 12) suggested a two-factor solution, which accounted for 65.7% of the variance. The experiential indicators loaded clearly onto factor 1, and the behavioural indicators loaded clearly onto factor 2 (see Table 5.4).

Table 5.4. Factor loadings of the ‘DBCI Engagement Scale’ in EFAs.

Scale Items	**Solution 1***			**Solution 2****		**Solution 3*****	
	Factor 1	**Factor 2**	**Factor 3**	**Factor 1**	**Factor 2**	**Factor 1**	**Factor 2**
1. Interest	**0.75**	0.14	0.25	**0.80**	0.26	**0.82**	0.28
2. Intrigue	**0.51**	0.09	0.11	**0.55**	0.09	**0.55**	0.18
3. Focus	**0.87**	0.28	0.09	**0.83**	0.02	**0.80**	0.27
4. Inattention (R)	0.25	**0.61**	0.14	N/A	N/A	N/A	N/A
5. Distraction (R)	0.21	**0.68**	0.06	N/A	N/A	N/A	N/A
6. Enjoyment	**0.66**	-0.35	0.43	**0.57**	0.31	**0.57**	0.23
7. Pleasure	0.31	-0.48	**0.56**	N/A	N/A	N/A	N/A
8. Annoyance (R)	**0.41**	0.23	0.25	N/A	N/A	N/A	N/A
9. Which of app’s components	0.16	0.01	**0.43**	0.15	**0.55**	N/A	N/A
10. How much time spent	0.10	0.03	**0.64**	0.09	**0.53**	N/A	N/A
11. Objective depth of use	N/A	N/A	N/A	N/A	N/A	0.37	**0.77**
12. Objective amount of use	N/A	N/A	N/A	N/A	N/A	0.18	**0.68**

Note. The symbol (R) indicates that values have been reverse scored prior to analysis. *EFA with oblique rotation, including items 1-10; ** EFA with oblique rotation, including items 1, 2, 3, 6, 9 and 10; *** EFA with oblique rotation, including items 1, 2, 3, 6, 11 and 12.

Solution 2 was selected for further analyses, as it contained only the self-reported items. Prior to this, a total scale score was calculated for each participant, with equal weight given to each of the retained self-reported items (i.e. items 1, 2, 3, 6, 9 and 10).

5.4.3.2 <u>Internal consistency reliability</u>

The internal consistency of the overall measure was α = .67, indicating questionable internal reliability [298]. The 'Experiential Engagement' subscale had an internal consistency of α = .78, while the 'Behavioural Engagement' subscale had an internal consistency of α = .45. Both subscales were significantly correlated with the measure overall ($r(145) = .90$, $p < .001$; $r(145) = .56$, $p < .001$, respectively). However, the subscales were not significantly correlated with each other ($r(145) = .15$, $p = .07$).

5.4.3.3 <u>Criterion validity</u>

Total scale scores were significantly correlated with objectively recorded 'depth of use', $r(145) = 0.32$, $p < .001$, and with objectively recorded 'amount of use', $r(145) = 0.33$, $p < .001$. Self-reported 'depth of use' was significantly correlated with objectively recorded 'depth of use', $r(145) = 0.51$, $p < .001$. Self-reported 'amount of use' was not significantly correlated with objectively recorded 'amount of use', $r(145) = 0.10$, $p = .23$.

5.4.3.4 Predictive validity

Results from the regression analyses are reported in Table 5.5. In an unadjusted model, total scale scores significantly predicted the variable 'subsequent login' (OR = 1.15, 95% CI = 1.05-1.27, p = .01).

Asking users about how engaging they thought the app was did not significantly predict the variable 'subsequent login' (OR = 1.28, 95% CI = 0.96-1.71, p = .10). Asking users about how much they liked the app was a significant predictor of whether or not they opened the app again (OR = 1.39, 95% CI = 1.05-1.83, p = .02).

Motivation to reduce alcohol consumption was not a significant predictor of total scale scores (B = 0.26, 95% CI = -0.05-0.57, p = .09), but being highly motivated to reduce alcohol (as compared with the other levels of motivation) was a significant predictor of the variable 'subsequent login' (OR = 5.40, 95% CI = 1.22-23.96, p = .03). In a model adjusting for motivation to reduce alcohol, the association between total scale scores and the variable 'subsequent login' remained significant (OR = 1.14, 95% CI = 1.03-1.27, p = .01). Being highly motivated to reduce alcohol was no longer a significant predictor of whether or not participants opened the app again (OR = 4.43, 95% CI = 0.97-20.29, p = .06).

When adjusting for motivation to reduce alcohol, the association between how engaging participants thought the app was and the variable 'subsequent login' remained non-significant (OR = 1.34, 95% CI = 0.98-1.84, p = .07). When adjusting for motivation to reduce alcohol, the association between how much participants liked the app and the variable ‘subsequent logins’ remained significant (OR = 1.38, 95% CI = 1.03-1.84, p = .03).

Table 5.5. Adjusted and unadjusted logistic regression models predicting the variable 'subsequent login'.

	OR (95% CI)	**OR_{adj} (95% CI)**
Predictive Validity		
Total scale scores	1.15 (1.05-1.27)*	1.14 (1.03-1.27)*
How engaging was the app?	1.28 (0.96-1.71)	1.34 (0.98-1.84)
How much did you like the app?	1.39 (1.05-1.83)*	1.38 (1.03-1.84)*
MTSS		
I don’t want to cut down on drinking alcohol (reference)	1	-
I think I should cut down on drinking alcohol but I don’t really want to	1.30 (0.37-4.52)	-
I want to cut down on drinking alcohol but I haven’t thought about when	2.48 (0.60-10.27)	-
I really want to cut down on drinking alcohol but I don’t know when I will	1.26 (0.29-5.42)	-
I want to cut down on drinking and hope to soon	0.96 (0.24-3.85)	-
I really want to cut down on drinking alcohol and intend to in the next 3 months	0.68 (0.12-3.78)	-
I really want to cut down on drinking alcohol and intend to in the next month	5.40 (1.22-23.96)*	-
Incremental Validity		
Model 1		
Objective amount of use	3.48 (1.59-7.61)**	-
Objective depth of use	0.91 (0.58-1.42)	-
Model 2		
Objective amount of use	2.88 (1.26-6.58)*	-
Objective depth of use	0.95 (0.60-1.50)	-
Interest	1.72 (1.03-2.85)*	-
Focus	0.82 (0.50-1.35)	-
Enjoyment	0.93 (0.61-1.40)	-
Intrigue	1.17 (0.78-1.76)	-

Note. OR_{adj} = ORs adjusted for the Motivation To Stop Scale; * p < .05; ** p < .01.

5.4.3.5 Incremental validity

Results from the regression analyses are reported in Table 5.5. A model including the automatically recorded indicators of engagement (i.e. items 11 and 12) accounted for 16% of variance in the variable 'subsequent login' (Model 1). A model including the automatically recorded indicators of engagement in addition to the experiential indicators (i.e. items 1, 2, 3 and 6) accounted for 21% of variance in the variable 'subsequent login' (Model 2). Of the experiential indicators, interest was the only independent, significant predictor of the variable 'subsequent login' (OR = 1.72, 95% CI = 1.03-2.85, p = .04).

5.4.3.6 Divergent validity

Total scale scores were significantly correlated with the first ("When using *Drink Less*, the way time passed seemed different from normal") but not the second ("When using *Drink Less*, I was not worried about what others may have been thinking about me") item tapping flow ($r(145) = 0.25$, $p < .01$; $r(145) = -0.01$, p = .95, respectively). The two items tapping flow were not significantly correlated with one another in this sample ($r(145) = -0.06$, p = .47).

5.5 Discussion

5.5.1 Summary of key findings

This study evaluated the psychometric properties of the 'DBCI Engagement Scale' in a sample of adult, UK-based, excessive drinkers who were willing to download and explore the *Drink Less* app in exchange for a financial reward. To ensure that participants with a broad range of engagement levels were recruited, participants in the present study were recruited via Prolific, an online, web-based platform designed specifically for recruiting and paying participants to complete research tasks.

First, although 266 participants were eligible to participate, only 147 (55%) participants completed the task. Secondly, results from a series of EFAs indicated that a two-factor solution provided the most appropriate fit, with the first factor labelled ‘Experiential Engagement’ and the second factor labelled ‘Behavioural Engagement’. Thirdly, the retained 6-item scale did not demonstrate adequate internal consistency reliability, or divergent validity. Although the overall measure demonstrated adequate criterion validity, self-reported ‘amount of use’ was not significantly correlated with objective ‘amount of use’. Fourthly, the overall measure of engagement was weakly associated with whether or not participants opened the app again. This association remained significant in a model adjusting for motivation to reduce alcohol. Motivation to reduce alcohol was strongly associated with engagement, but was not predictive of participants’ total engagement scores. Finally, asking participants about how much they liked the app (but not how engaging they thought the app was) was weakly associated with the variable ‘subsequent login’.

In contrast to the first attempt at evaluating the ‘DBCI Engagement Scale’ (reported in Chapter 4), results from the present study indicated that a two-factor solution provided the best fit of the data. The ‘Experiential Engagement’ and ‘Behavioural Engagement’ subscales were not significantly correlated with each other. Hence, the results from the present study lend support to the argument put forward in Chapters 2 and 4, namely that users can spend time on a DBCI without necessarily being interested in or paying attention to its content (and vice versa). Moreover, the finding that motivation to reduce alcohol was not predictive of initial experiential and behavioural engagement suggests that the state of engagement with a DBCI may be distinct from motivation to change the target behaviour.

With regards to the scale's criterion validity, the lack of a significant correlation between self-reported and objectively recorded 'amount of use' in this sample can potentially be explained in light of findings from the analysis of the scale's divergent validity. In line with the first evaluation study, the present study did not provide evidence that the 'DBCI Engagement Scale' diverges from the 'Flow State Scale'. It appears that there may be conceptual overlap between the state of engagement with DBCIs and the subdimension of flow that is labelled 'losing track of time'. As the proposed definition of engagement was in part developed based on the concept of flow, these findings are not unexpected. However, this serves as a plausible explanation for why participants' estimates of their amount of use were not significantly correlated with their objective amount of use; they may have lost track of time when engaging with the *Drink Less* app. In the incremental validity analyses, objective amount of use was found to be strongly related to the variable 'subsequent login'. Hence, further work is required to explore whether it is more useful to consider objective or subjective amount of time spent on a DBCI going forward. Moreover, it may be more fruitful to test the scale's divergent validity using a more conceptually distinct measure in the future.

In contrast to the first evaluation study, participants' overall scores on the 'DBCI Engagement Scale' were weakly associated with future behavioural engagement in both unadjusted and adjusted analyses. However, it is difficult to theorise about what a clinically meaningful effect would look like in this context, as further research linking initial and future engagement to intervention effectiveness is also required. The finding that a 1-point increase on the standardised ‘DBCI Engagement Scale’ is associated with 1.15 times greater odds of opening the app again might be found to be a meaningful effect.

One possible explanation for the finding that initial engagement predicted future engagement when adjusting for baseline motivation to change, is that more intensive engagement with the *Drink Less* app led to an increase in participants’ motivation to reduce their drinking. This might in turn have made them more prone to return to the app. Alternatively, more intensive engagement during the first login session might have made users’ memory of the app more salient, which made them more likely to remember to return to the app. As the short measure of how much users liked the app was also found to be predictive of subsequent engagement, it is possible that not only salience of the app, but a salient memory that one liked the app, is important for future engagement. It is unclear why the first, but not the second short measure of engagement was found to have predictive power. This could potentially be explained by the word ‘liking’ being easier to interpret than the word ‘engaging’. The potential mechanisms underlying the relationship between initial experiential and behavioural engagement, and future behavioural engagement, should be explored further using experience sampling techniques, which involves repeated measurements of participants’ psychological states and behaviours in real-time [226], in the first few hours following initial app engagement.

5.5.2 Limitations

This study has a few important limitations. Although a sample with a broader level of engagement levels was recruited into the present study (as compared with the study reported in Chapter 4), the desired sample size of 250 participants was not reached. Although it has been specified in the literature that the participant-to-item ratio is key in determining the minimum necessary sample size for conducting factor analyses, findings from simulation studies indicate that other factors, including the number of items per factor and the level of communality between items also influence sample size requirements [313]. Given the small number of items per factor and the wide-ranging level of item communality in the present study, the two-factor solution should be interpreted with caution and needs to be replicated in a larger sample in future research.

Longitudinal studies conducted via Prolific involve an initial screening study, with eligible users subsequently being invited to complete the actual study. Such studies tend to observe attrition rates of approximately 20-25%, and not 45% [314]. It is therefore likely that there were systematic differences between the eligible participants who completed the task and those who did not. Potential explanations are that participants experienced technical issues when trying to complete the task (as evidenced by some participants' responses to the reminder messages) or that the small financial reward was not seen as worth the effort. Indeed, a study assessing the demographic and psychological characteristics of participants who regularly complete research tasks via Amazon's Mechanical Turk online platform (which is similar to Prolific) found that 61% of surveyed participants (N = 1,000) reported that earning money was a key motivator of participation [315].

5.5.3 Conclusion

Behavioural and experiential indicators may constitute different dimensions of engagement. Initial engagement, both behavioural and experiential, predicted future behavioural engagement, and this held true when adjusting for baseline motivation to change. Due to not achieving the desired sample size, these findings require replication in a larger sample before drawing firm conclusions regarding the scale's factor structure and its relationship with key outcome variables.

5.5.4 Next steps

The remaining empirical studies in this thesis focused on the identification of factors that influence engagement with alcohol reduction apps. The next study investigated the design features considered to be most important for engagement with apps for alcohol reduction through the use of a novel ranking task paradigm (Study 5, reported in Chapter 6). In line with a user-centred design approach, Study 5 aimed to elicit potential app users' needs and preferences, with a view to using this information to inform the design of new, and the modification of existing, apps for alcohol reduction.

6 CHAPTER 6 – Engagement features judged by excessive drinkers as most important to include in smartphone applications for alcohol reduction: A mixed-methods study (Study 5)

6.1 Abstract

Objective: Engagement with apps for alcohol reduction is necessary for their effectiveness. This study explored 1) the design features that are ranked as most important for engagement by excessive drinkers and 2) why particular design features are judged to be more important for engagement than others.

Methods: Two studies were conducted in parallel. The first was a focus group study with adult excessive drinkers, interested in reducing alcohol consumption using an app (N_{groups} = 3). Participants individually ranked their top 10 features from a pre-specified list and subsequently discussed their rankings. The second was an online study with a new sample (*N* = 132). Rankings were analysed using the intraclass correlation coefficient (ICC) to assess level of agreement between raters for each study. Qualitative data were analysed using inductive thematic analysis.

Results: There was low agreement between participants in their rankings, both in the focus groups (ICC = 0.15, 95% CI = 0.03-0.38) and the online sample (ICC = 0.11, 95% CI = 0.06-0.23). 'Personalisation', 'control features' and 'interactive features' were most highly ranked in the focus groups. These were expected to elicit a sense of benefit and usefulness, adaptability, provide motivational support or spark users' interest. Results from the online study partly corroborated these findings.

Conclusion: There was little agreement between participants, but on average, the features judged to be most important for inclusion in smartphone apps for alcohol reduction were personalisation, interactive features and control features.

6.2 Introduction

Although existing alcohol reduction apps have involved users in the design process [97,316,317], thus increasing their engagement potential, the benefits of such user-centred design activities may be limited by involving only a small number of potential users in the process. Although this allows researchers and designers to gain an in-depth understanding of users' needs, insights from a small number of highly motivated participants who are willing to take part in design sessions may not generalise to other target users. For example, although community drug and alcohol service users were involved in the design of *DIAMOND*, a web-based alcohol intervention, few new patients recruited from the same service were willing to be randomised in a feasibility trial, mainly due to expressing a strong preference for face-to-face treatment [318].

The present study used a mixed-methods approach, combining focus group methodology with an online study, to identify engagement features judged by excessive drinkers as most important to include in smartphone apps for alcohol reduction. In-depth focus group discussions were conducted in a small sample, in parallel with an online study with a larger sample of excessive drinkers, to address the following research questions:

1. What engagement features are ranked most highly by potential users of alcohol reduction apps?

2. What reasons do potential users give for judging particular features to be more important for engagement than others?

6.3 Methods

6.3.1 Study design

Two parallel studies were conducted. The first was a focus group study and the second was an online study. As both methods have a number of well-known strengths and weaknesses, data sources were triangulated to address the same research questions.

Focus groups are useful for gaining an in-depth understanding of participants' experiences, beliefs and motivations, and are particularly suitable when the interaction between participants is expected to yield additional insight into the topic of interest [319]. Hearing about others' experiences and views may stimulate discussion and allow participants to elaborate on ideas mentioned by other group members [320]. However, a key weakness is that focus groups may inhibit the expression of controversial opinions due to social conformity, thus restricting the understanding of the diversity of users' needs and preferences [320].

Research conducted online benefits from being able to reach larger, geographically diverse samples. Hence, results from online surveys are more likely to generalise to other members of the target population than findings from focus groups. Despite these strengths, online surveys that require cognitive effort may suffer from 'satisficing', meaning that respondents simply provide a satisfactory answer or randomly choose among response options [321,322].

6.3.1.1 Participants

6.3.1.1.1 Focus groups

Drinkers were eligible to participate in one of the focus groups if they i) were aged ≥ 18 years, ii) lived in in or near London (UK), iii) reported an AUDIT score of ≥ 8, indicating excessive alcohol consumption [77], iv) owned an Android or iOS smartphone with internet access and v) were interested in using a smartphone app to reduce their drinking and vi) had previously used a health or fitness app. It was expected that participants with prior experience of using a health or fitness app would be able to more vividly imagine whether or not a particular feature would be important for engagement and hence generate more valid data.

Participants were recruited online through Gumtree (www.gumtree.com) and Call for Participants (www.callforparticipants.com) in addition to posters placed on central London university campuses. The recruitment materials stated that drinkers were invited to the laboratory to contribute to a focus group discussion with other participants about how to design engaging smartphone apps for alcohol reduction (see Appendix 9).

Of 48 participants who completed the screening questionnaire, 29 were eligible to take part. Thirteen participants did not respond to any further study communications. Six participants cancelled prior to taking part. One participant failed to arrive on time. In total, nine participants took part in one of three focus groups, with three participants in each group (see Figure 6.1). The average age of participants was 30.0 years (*SD* = 10.1), 77.8% were female and 66.7% had a non-manual occupation. Participants had an average AUDIT score of 13.6 (*SD* = 3.1), indicating excessive alcohol consumption (see Table 6.1).

6.3.1.1.2 Online sample

A new sample of drinkers was recruited into the online study. Participants were eligible if they met the inclusion criteria outlined above with the exception of ii) and vi). Instead, participants had to reside in the UK and did not need prior experience of using a health or fitness app. As the purpose was to explore generalisability, the online sampling was less restrictive. Eligible participants who did not pass a multiple-choice attention check at the end of the ranking task (i.e. “What is a professional support feature?”) were excluded from the analysis.

Participants were recruited online through Prolific (www.prolific.ac). The recruitment materials invited drinkers to familiarise themselves with sixteen different engagement features and rank their top 10 choices based on their likelihood of promoting engagement with apps for alcohol reduction.

Of 400 participants who completed the screening questionnaire, 181 were invited to complete the ranking task. Of these, 148 participants completed the ranking task, with 132 participants included in the analytical sample (see Figure 6.1). Just under half of the included participants were female (49.2%), 34.1% were aged 35-44 years, 13.6% had a manual occupation and 70.5% had a non-manual occupation. Participants had an average AUDIT score of 16.1 (*SD* = 6.7), indicating excessive alcohol consumption (see Table 6.1).

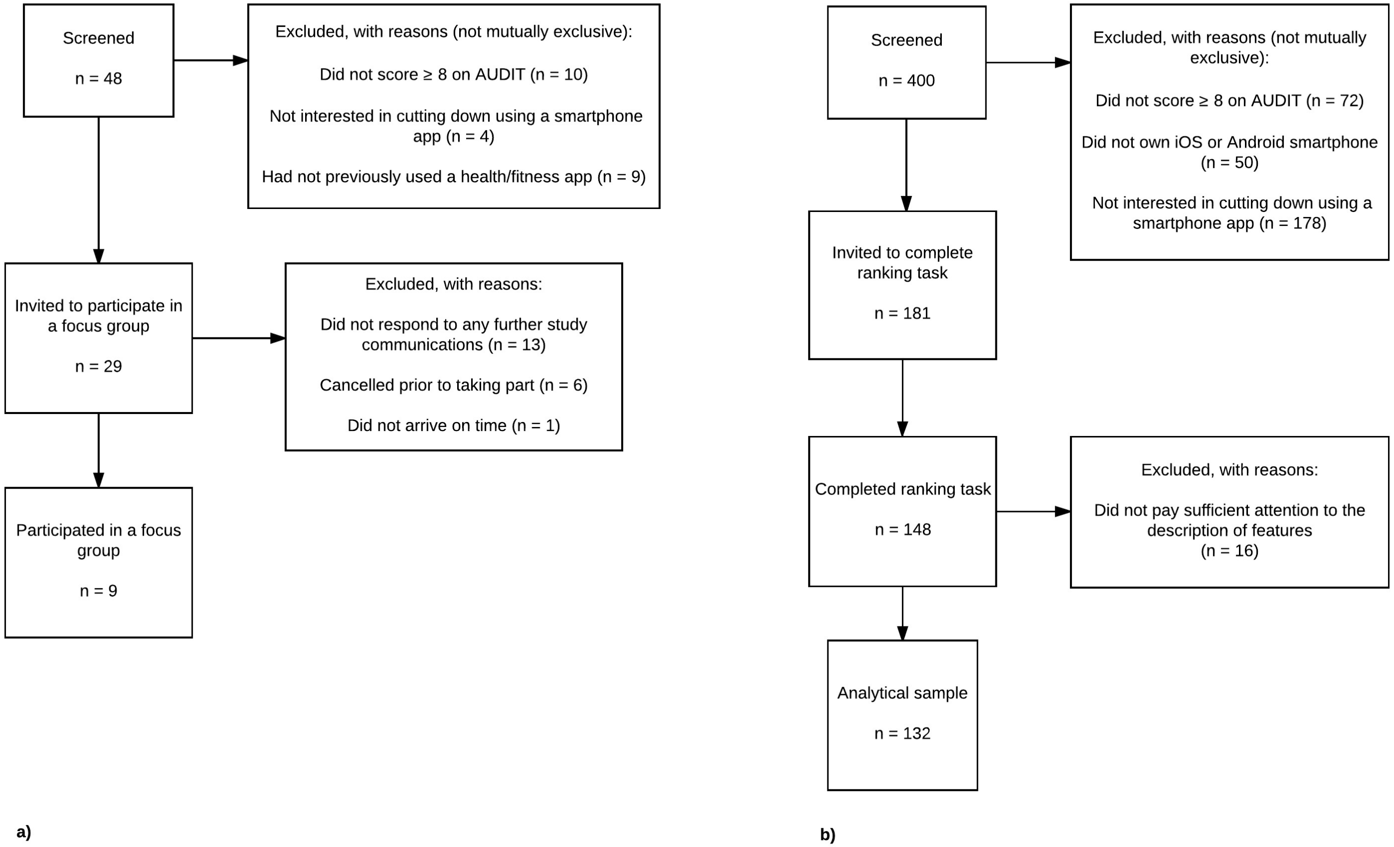


Figure 6.1. Participant flow charts for a) the focus group study, and b) the online sample.

Table 6.1. Participants' demographic and drinking characteristics.

Demographic and drinking characteristics	***Focus groups,* n (%)**	***Online sample,* n (%)**
Gender		
Women	7 (77.8%)	65 (49.2%)
Men	2 (22.2%)	67 (50.8%)
Age (years)		
18-24	4 (44.4%)	14 (10.6%)
25-34	3 (33.3%)	32 (24.2%)
35-44	0 (0%)	45 (34.1%)
45-54	2 (22.2%)	28 (21.2%)
55-64	0 (0%)	9 (6.8%)
65+	0 (0%)	4 (3.0%)
Type of work		
Manual	0 (0%)	18 (13.6%)
Non-manual	6 (66.7%)	93 (70.5%)
Other	3 (33.3%)	21 (15.9%)
AUDIT, mean (*SD*)	13.6 (3.1)	16.1 (6.7)
MTSS*		
1. I don't want to cut down on drinking alcohol	1 (11.1%)	8 (6.1%)
2. I think I should cut down on drinking alcohol but I don't really want to	1 (11.1%)	42 (31.8%)
3. I want to cut down but haven't thought about when	4 (44.4%)	16 (12.1%)
4. I really want to cut down but I don't know when I will	0 (0%)	10 (7.6%)
5. I want to cut down and hope to soon	1 (11.1%)	18 (13.6%)
6. I really want to cut down and intend to in the next 3 months	0 (0%)	10 (7.6%)
7. I really want to cut down and intend to in the next month	2 (22.2%)	28 (21.2%)

Note. * MTSS = Motivation to Stop Scale.

6.3.1.2 Measures

Data were collected on: 1) age; 2) gender; 3) type of work (i.e. manual, non-manual, other); 4) alcohol consumption, measured using the AUDIT; 5) interest in using a smartphone app to help cut down on alcohol (yes vs. no); and 6) motivation to cut down on drinking alcohol, measured using the MTSS.

6.3.1.3 Materials

Sixteen different engagement features, derived from the systematic review reported in Chapter 1 [281], were used as stimuli (see Table 6.2). Feature descriptions were piloted and refined based on feedback from four independent researchers and five non-expert app users, recruited from the author's networks. Engagement features that have previously been found to be difficult for participants to describe verbally (e.g. aesthetics, ease of use, message tone) were not included. An experimental study design was expected to generate more valid data about how such abstract features influence engagement [287].

Table 6.2. Engagement features used in the ranking task.

Engagement features	*Descriptions and examples*
Challenge features	Features that allow you to compete against yourself or against other users, such as your friends. The app might, for example, encourage you to drink one unit less than your friends.
Control features	Features that allow you to make choices about how to use the app. The app might, for example, allow you to choose between a few different target goals instead of having one fixed option.
Action plans to use the app	A feature that encourages you to make a plan to use the app. An example might be to make a plan to open the app as soon as you have finished your breakfast every morning.
Setting a goal to use the app	A feature that encourages you to set a goal to use the app. For example, you might be able to set a goal to use the app once a day for two weeks.
Monitoring use of the app	A feature that allows you to record your use of the app. For example, the app might allow you to manually enter how much time you have spent on it, or it might record it automatically for you.
Feedback on use of the app	A feature that allows you to view your use of the app. For example, the app might show you how many times you have opened it on each day of the week.
Credibility features	Features that make you feel that you can trust the app. For example, the app might have a clear privacy policy, be endorsed by a trusted organisation, or be free from adverts.
Guidance features	Features that explain how to use the app. This might, for example, include video tutorials about how the app's different features work.

Table 6.2. *Continued.*

Engagement features	***Descriptions and examples***
Interactive features	Features that allow and respond to input from the user. This might, for example, include a game or a knowledge quiz. The direct opposite would be a static app that does not allow you to enter any information or click into any of its features, much like this piece of text!
Novelty features	Features that ensure that you see or learn something new every time you open the app. This might, for example, include daily content updates (e.g. a daily fact about alcohol or a daily motivational quote).
Narrative features	The presence of a storyline. For example, the app might be set up as a game or film with a plot, where you are the main character. This might include the presence of an avatar (i.e. a virtual figure that represents you).
Personalisation	Tailoring of content according to information about you (driven by the app) or customisation of the app so that it looks or acts the way you prefer (driven by you). For example, the app might tailor its content based on information you give to it (e.g. about your age, gender, level of alcohol consumption) or you might be able to change the colour and font.
Professional support features	Features that enable you to have remote contact with a healthcare professional (e.g. the opportunity to chat to a nurse or a psychologist via the app).
Social support features	Features that allow you to connect with other app users. This might, for example, include an online discussion forum or a peer-to-peer instant messenger (e.g. a 'buddy system').
Reminders to use the app	Regular push notifications or text messages that remind you to use the app.
Rewards for using the app	Being rewarded for using the app. You might, for example, receive a congratulatory message or a virtual badge/coin after having opened the app for seven days in a row.

6.3.1.4 ***Procedure***

Interested participants read the information sheet describing the study. They subsequently provided informed consent via an online screening questionnaire, which also assessed study eligibility and collected descriptive data. The screening questionnaire was hosted by Qualtrics survey software [266].

6.3.1.4.1 Focus groups

The focus groups were conducted at UCL. Sessions lasted approximately 2 hours. Participants received a £20 gift voucher as compensation for their time. Sessions were facilitated by the author with support from a second researcher.

6.3.1.4.1.1 Individual activity

An individual activity was first conducted to allow participants to familiarise themselves with the engagement features and to elicit their attitudes to the features. The term ‘engagement’ was defined as a behaviour (e.g. how often you use the app, how much time you spend on it) and as an experience (e.g. how interested you are in the app, how much attention you pay to it, how much you enjoy using it) [281].

Participants were each given a folder with post-its. Each of the 16 engagement features was described on a separate post-it, accompanied by an illustrative example. Participants were also encouraged to think of their own examples. They were asked to rank their top 10 choices without consulting the other participants, and were subsequently asked to place the post-its with their selected features on a whiteboard, thus sharing their rankings with the group.

6.3.1.4.1.2 Group discussion

Participants subsequently convened to discuss their rankings. A semi-structured topic guide was used to steer the discussion (see Appendix 10). To gain a better understanding of why particular features were perceived as more important for engagement than others, participants were prompted to discuss the reasons for their rankings (e.g. “Can you tell me a bit more about why you ranked [insert feature here] highly?”).

6.3.1.4.2 Online sample

Eligible participants were invited to complete the online ranking task in their own time on a personal computer, tablet or smartphone. The ranking task lasted for approximately 10 minutes and was hosted by Qualtrics survey software. Participants were paid £0.85 as compensation for their time. Participants were asked to complete the same ranking task as the focus group participants. At the end of the ranking task, participants were asked to respond to a multiple-choice attention check (described above). To gain a better understanding of why particular features were ranked more highly than others, participants were asked to respond to a free-text question about why they believed that their top choice would be important for engagement.

6.3.2 Data analysis

6.3.2.1 <u>Focus groups</u>

Participants assigned a unique score from 1-10 to their top ten engagement features, with 1 representing their top choice. The remaining 6 features were assigned a rank of 11, as the distance between these features was not expected to be meaningful. To assess the level of agreement between participants, the intraclass correlation coefficient (ICC) was estimated by means of a single measurement, absolute agreement, two-way, mixed effects model. To assess whether some of the engagement features were, on average, ranked more highly than others, rankings were reverse scored (to aid interpretation) and descriptive statistics were calculated.

Sessions were audio-recorded, transcribed verbatim and analysed using inductive thematic analysis. To inform the analysis, an interpretivist theoretical framework was used, based on the premise that the 'lived experience' of the individual can be captured through discussion between the researcher and participant [250]. The thematic analysis was conducted in six phases: i) gaining familiarity with the data, ii) generating initial codes, iii) searching for themes, iv) reviewing themes, v) defining and naming themes, and vi) producing the report [267]. Data were independently coded by the author and a second, independent researcher. New inductive codes were labelled as they were identified during the coding process. Data were sometimes assigned to multiple codes. All codes that included data relating to the research questions were recorded. The author reviewed the codes one by one, ordering the findings systematically under headings. The ordered data were reviewed and revised in discussion with the second researcher and were subsequently organised into themes. Disagreements were resolved through discussion. Agreement on the final themes was reached through discussion between all members of the supervisory team.

6.3.2.2 <u>Online sample</u>

Participants who provided incorrect responses to the ‘attention check’ were excluded from the analysis, as incorrect responses were interpreted to suggest that participants had not paid sufficient attention to the task to provide valid data [322]. A single measurement, absolute agreement, two-way, mixed effects model was fitted to estimate the ICC. Rankings were reverse scored and descriptive statistics were calculated.

Responses to the free-text question about why participants believed that their top choice would be important for engagement were analysed using inductive thematic analysis (described in Section 6.3.2.1).

6.3.3 Ethical approval

Ethical approval was granted by UCL’s Departmental Research Ethics Committee (UCLIC/ 1213/015). Personal identifiers were removed and data were stored securely.

6.4 Results

6.4.1 Engagement features ranked most highly by potential users of alcohol reduction apps

6.4.1.1 <u>Focus groups</u>

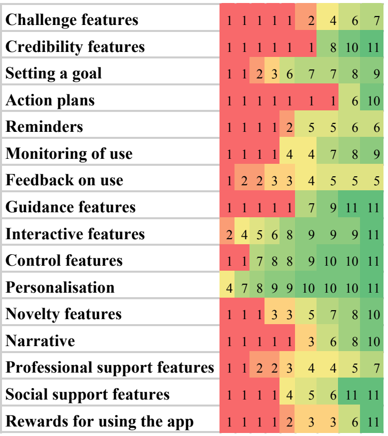
There was positive but low agreement between participants (ICC = 0.15, 95% CI = 0.03-0.38; see Figure 5.2). On average, participants ranked personalisation (*M* = 8.67, *SD* = 2.12), control features (*M* = 7.22, *SD* = 3.73) and interactive features (*M* = 7.00, *SD* = 2.92) most highly. Action plans (*M* = 2.56, *SD* = 3.24) and challenge features (*M* = 2.67, *SD* = 2.40) were judged to be least important for engagement (see Table 6.3 and Figure 6.2).

6.4.1.2 Online sample

There was positive but low agreement between participants (ICC = 0.11, 95% CI = 0.06-0.23; see Figure 5.2). On average, participants ranked personalisation (*M* = 6.74, *SD* = 3.18), setting a goal to use the app (*M* = 5.97, *SD* = 3.66) and challenge features (*M* = 5.56, *SD* = 3.93) most highly. Narrative features (*M* = 2.26, *SD* = 2.53) and feedback on use of the app (*M* = 2.68, *SD* = 2.33) were judged to be least important for engagement (see Table 6.3 and Figure 6.2).

Table 6.3. Mean rankings of the 16 engagement features in a) the focus groups (N = 9) and b) the online sample (N = 132).

a) Focus groups		***b) Online sample***	
Engagement features	***Mean (SD)***	***Engagement features***	***Mean (SD)***
1. Personalisation	8.67 (*2.12*)	**1. Personalisation**	6.74 (*3.18*)
2. Control features	7.22 (*3.73*)	**2. Setting a goal to use the app**	5.97 (*3.66*)
3. Interactive features	7.00 (*2.92*)	**3. Challenge features**	5.56 (*3.93*)
4. Setting a goal to use the app	4.89 (*3.14*)	**4. Interactive features**	5.43 (*3.39*)
5. Guidance features	4.78 (*4.63*)	**5. Control features**	5.41 (*3.40*)
6. Social support features	4.56 (*4.13*)	**6. Credibility features**	4.86 (*3.99*)
7. Novelty features	4.33 (*3.35*)	**7. Rewards for using the app**	4.70 (*3.49*)
8. Monitoring of use	4.00 (*3.28*)	**8. Professional support features**	4.36 (*3.55*)
9. Credibility features	3.89 (*4.40*)	**9. Reminders**	4.27 (*3.20*)
10. Narrative features	3.56 (*3.54*)	**10. Social support features**	3.82 (*3.31*)
11. Feedback on use	3.33 (*1.50*)	**11. Action plans**	3.98 (*3.19*)
12. Professional support features	3.22 (*1.99*)	**12. Guidance features**	3.74 (*3.31*)
13. Rewards for using the app	3.22 (*3.35*)	**13. Novelty features**	3.66 (*3.16*)
14. Reminders	3.11 (*2.32*)	**14. Monitoring of use**	3.56 (*3.02*)
15. Challenge features	2.67 (*2.40*)	**15. Feedback on use**	2.68 (*2.33*)
16. Action plans	2.56 (*3.24*)	**16. Narrative features**	2.26 (*2.53*)



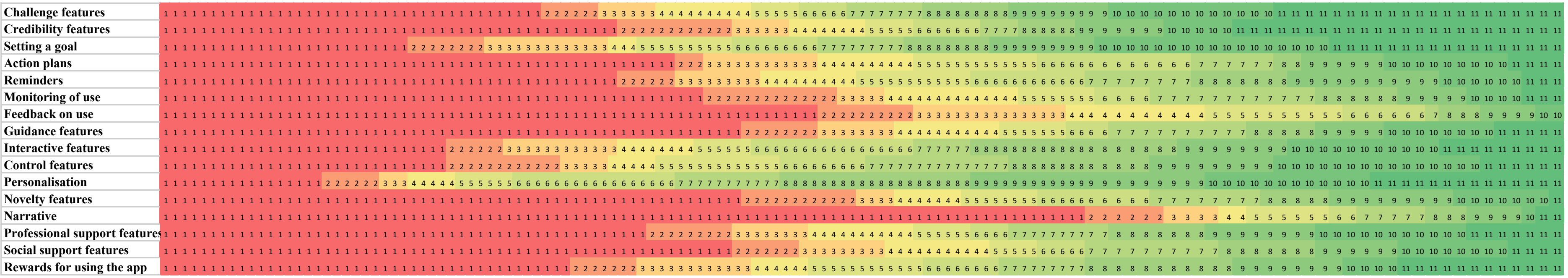


Figure 6.2. Heat maps of rankings in the focus groups (top) and in the online sample (bottom). Red, orange and yellow boxes indicate low rankings. Green boxes indicate high rankings.

6.4.2 Judgments as to why particular features are expected to be more important for engagement than others

Six themes were generated: ‘lack of trust and guidance as initial barriers’, ‘motivational support’, ‘benefit and usefulness’, ‘adaptability’, ‘sparking users’ interest’ and ‘relatedness’. Two subthemes were developed in relation to the final theme, which were labelled ‘perceived social stigma’ and ‘fear of social comparison’ (see Table 6.4). Additional quotations can be found in Appendix 11.

Table 6.4. Summary of themes and subthemes identified in a) the focus groups and b) the online sample.

Themes	***Description***	***a) Identified in focus groups***	***b) Identified in online sample***
1. Lack of trust and guidance as initial barriers	Features that inculcate feelings of trust and ensure that the user can use the app comfortably (e.g. credibility features, guidance features) were considered to be more important for initial uptake than for continued engagement.	√	√
2. Motivational support	Features that support users' motivation to engage with the app or to cut down on drinking (e.g. control features, rewards, setting a goal to use the app, challenge features, message tone) were expected to encourage engagement, particularly if they promote users' independence.	√	√
3. Benefit and usefulness	Features that make users feel that they are gaining something over and above *status quo* (e.g. personalisation, interactive features, novelty features, rewards) were expected to prompt engagement, particularly if they have utility 'in real life'.	√	√
4. Adaptability	Features that allow the app to adapt its content according to the user's level of progress or to intervene in the right moment (e.g. personalisation, interactive features, reminders) were expected to persuade the user and hence, promote engagement.	√	√
5. Sparking users' interest	Features that grab users' interest or provide a means of entertainment (e.g. narrative features, social support features, challenge features, interactive features, novelty features) were expected to prompt engagement.	√	√
6. Relatedness	Features that allow the user to connect with others who are in the same situation (e.g. social support features) were expected to promote engagement.	√	√
i. Perceived social stigma	Features that trigger app use in front of family and friends or that connect users with close others (e.g. social support features, challenge features) were expected by some participants to elicit feelings of embarrassment and hence, lead to disengagement.	√	
ii. Fear of social comparison	Features that encourage users to compete against friends or strangers (e.g. challenge features) were expected by some participants to be demoralising.	√	

6.4.2.1 Lack of trust and guidance as initial barriers

Although participants expected the presence of credibility features to be necessary to decide whether or not to engage with the app in the first place (as such features would inculcate feelings of trust), they did not believe that credibility features would promote further engagement after having made an initial decision to download an app.

> *"…it wouldn't increase my engagement behaviour. It would just be the barrier, and make sure that I would actually use it, rather than frequently use it." – P2, focus group*

Similarly, the presence of guidance features was expected to aid initial app navigation, but was not expected to prompt continued engagement. If guidance was provided again later, this was expected to be annoying, as participants believed that they would be capable of using the app without any further support.

> *"Just at the beginning of the app, when you've downloaded it and you're using it for the first time, it should tell you what to do. But not every time. You don't need guidance how to use it and where things are, because I think it would just be annoying..." – P3, focus group*

6.4.2.2 Motivational support

Participants expected that features that provide motivational support would be important for engagement (e.g. control features, rewards, setting a goal to use the app, challenge features). This included features that support independent decision-making by, for example, allowing users to make choices about how to use the app (e.g. control features). Participants expected to feel more motivated to work towards achieving goals that they had set for themselves.

"I feel that if you decide to carry out a task, you need to be in control of it, because ultimately, that's your goal that you're setting, and you want to have a sense of ownership or control of whatever you want to achieve. You feel more responsible for how you carry out your goals." – P2, focus group

"The more I would be able to manipulate the app to be and do what I wanted or needed, for my own circumstances, the more likely I am to use it." – P16, online sample

The app's 'tone of voice' or the way in which feedback was framed was expected to influence engagement. For example, feedback on drinking patterns framed in a positive manner (i.e. gain- rather than loss-framed) was expected to enhance users' beliefs about their ability to cut down on alcohol, and hence motivate engagement with the app.

"...so that you don't feel discouraged when you drink too much, and then you decide that, you know what, I'm just going to ignore the app and shut it off…" – P8, focus group

Participants believed that setting a goal to use the app or the receipt of rewards would motivate them to return to the app. For example, virtual rewards (e.g. badges, points) were expected to automatically encourage engagement.

“It would encourage me to open the app on a daily basis.” – P37, online sample

“…even if it doesn’t have practical meaning, it still works, because it’s an incentive, and it tricks your brain to thinking that you’re earning…” – P3, focus group

Participants who ranked challenge features highly believed that competing against friends or other app users would help pushing oneself to achieve one’s targets, thus providing an important source of motivation to cut down on drinking.

“Personally, I feel if you have a community that challenges and pushes each other it encourages you to push yourself…” – P47, online sample

6.4.2.3 Benefit and usefulness

Participants believed that features that make users feel that they are gaining something over and above what they already knew or felt before downloading the app would be important for engagement (e.g. personalisation, interactive features, novelty features, rewards). For example, rewards that had utility ‘in real life’ or within the app itself (e.g. unlocking novel features, grocery store vouchers) were thought to be more likely to prompt engagement due to their real-world usefulness.

"Well, both of them are kind of: “Well done for doing this”, they’re both a reward, they both make you feel a bit better. But

a badge, it's a cool fact, but it's not the same as having vouchers, where you can go and treat yourself to something you want." – P6, focus group

Maintaining a balance between the amount of effort on the part of the user (e.g. inputting vast amounts of information) and the rewards or outputs received from the app was expected to be crucial for engagement. Participants believed that they would engage with the app only if they felt that they were getting something meaningful back, such as learning something new about alcohol or about themselves (e.g. through personalised feedback). They also expected that they would feel more warmly towards apps that maintained a two-way flow of communication between user and app (i.e. 'reciprocal interactivity').

"You've got to keep putting stuff in, but it's like, when am I going to get something out of it?" – P5, focus group

Participants who did not rank narrative features, action plans or goal setting to use the app highly believed that such features would distract from the main task of reducing alcohol consumption or be more effortful than rewarding.

"Well, surely the other features will make you want to use the app anyway." – P6, focus group

6.4.2.4 Adaptability

Participants expected that features that make users feel that the app adapts itself to their level of progress or intervenes in the right moment (e.g. personalisation, interactive features, reminders) would promote engagement due to inculcating the belief that the app is speaking directly to the user. Highly personalised and context-sensitive information was expected to be more persuasive than generic advice about how to drink less.

“If it’s personal to me, you just get a sense of uniqueness, and you’re like, yes, this is the best way for me to go, based on how I am right now…” – P2, focus group

“Every person is an individual, so I would have more faith in the app if it felt more tailored to my personal needs.” – P34, online sample

Participants also expected that features that allow the app to intervene either in the right moment or pre-emptively, ‘before it is too late’, would promote engagement. For example, participants who identified as heavy drinkers expected that professional support features would encourage engagement in ‘times of crisis’.

“It would help in times of crisis to be able to be in touch with a professional, or if I needed to ask health questions related to alcoholism.” – P51, online sample

However, participants who did not identify as having a problem with alcohol did not expect professional support features to encourage engagement.

"I think if I found that I had an issue with alcohol, maybe…" – P9, focus group

6.4.2.5 Sparking users' interest

Participants expected that the presence of features that grab users' attention or provide a means of entertainment (e.g. interactive features, narrative features, challenge features, social support features, novelty features) would prevent boredom and hence encourage users to return to the app. The hedonistic aspect of engagement was evident in participants' accounts, emphasising that some features are expected to be important for engagement only because they make the app more fun and enjoyable to use.

> *"An app without any interactivity would get boring very quickly, and I would probably forget about it or delete it after a while." – P72, online sample*

> *"I do think that you need to keep people slightly entertained..." – P9, focus group*

Participants who ranked social support features highly believed that features that connect the user with others would draw their attention to the app and hence, promote engagement with other features.

> *"If you saw a message from such and such, you might be more inclined to log on and respond to them. While you're on the app, you might use other features on it." – P6, focus group*

6.4.2.6 <u>Relatedness</u>

Participants who ranked social support features highly expected that such features would facilitate the receipt of non-judgmental support from other users and hence, foster a sense of relatedness.

> *"Being able to exchange feedback with strangers with the same goal could be supportive but non-judgemental as you will probably not know the other users." – P66, online sample*

6.4.2.6.1 Perceived social stigma

Participants who did not rank social support or challenge features highly imagined that features that trigger app use in front of family or friends or that connect users with others through the app would evoke feelings of embarrassment or worry that others may think that they have a problem with alcohol.

> *"…I wouldn't want something like: "Oh, why have you got that app?"" – P5, focus group*

6.4.2.6.2 Fear of social comparison

Participants who did not rank social support or challenge features highly also pointed out that such features may have a negative effect on motivation to change due to eliciting fear of failure or worry that others are progressing quicker than oneself.

"…somebody would always do better than me, performing better on the app than me, so I'd be engaging with people who are doing better than me on the app, which might be a bit demoralising…" – P4, focus group

6.5 Discussion

6.5.1 Summary of key findings

This mixed-methods study found that there was low agreement between participants concerning the importance of particular engagement features, both in the focus groups and in the online sample. In general, features judged to be most important for inclusion in smartphone apps for alcohol reduction were personalisation, control features and interactive features. These features were expected to foster a sense of benefit and usefulness, adaptability, provide motivational support or spark users' interest. Social support features and challenge features were ranked highly by a subset of participants as they were expected to foster relatedness and provide motivational support. However, another subset of potential users did not rank such features highly as they were expected to elicit social stigma or social comparison.

These findings lend support to and extend the results of prior research. First, there is previous support for the finding that personalisation is expected to promote engagement with alcohol reduction apps by inculcating the belief that the app is speaking directly to the user. Previous results have been consistent across types of study, including a formal expert consensus study [98] and a qualitative study with potential users [287]. This finding can be explained by the Elaboration Likelihood Model of Persuasion [63] and the Persuasive Systems

Design Model [61], which posit that messages tailored to users' needs and interests have greater potential for deep processing. These findings highlight two additional mechanisms through which personalisation may promote engagement. First, personalisation may help to foster a sense of benefit and usefulness. For example, encouraging users to return to the app to learn more about themselves by offering highly personalised suggestions may prevent users from feeling that they are inputting data without getting anything back. Secondly, personalisation may help to foster a sense of adaptability by supporting both user-led and reactive use. For example, participants imagined that they would engage more with apps that keep up-to-date with their progress and push relevant messages to users 'just-in-time'.

Secondly, previous research has emphasised the importance of features that support and enhance users' motivation [103,145,248]. Participants in the present study highlighted that they would be more motivated to achieve goals that they had set for themselves (i.e. the need for autonomy), suggesting that apps that provide autonomy-support in the form of 'control features' may be more conducive to longer-term engagement than those that do not [323]. However, participants in the present study also expected the receipt of rewards – which have previously been found to undermine participants' autonomy [324] – to help them engage with the app, begs the question as to what sources of motivation are most supportive of engagement.

Thirdly, these results suggest that users may continue to engage with alcohol reduction apps only if they are regularly provided with information or features that pique their interest. Although few studies in the alcohol domain have highlighted the importance of preventing boredom, this is not a novel idea in the digital gaming and technology literature [45,108]. It has been argued that users have 'non-instrumental' needs (i.e. needs that do not serve as a means to achieve a particular aim), such as the need for stimulation or enjoyment [56,57]. The presence of features that address these non-instrumental needs is expected to give rise to a positive user experience and hence encourage technology engagement [57]. It has also been suggested that it may be particularly important to sustain users' interest in the technology when they have deviated from their goals [325]. The possibility of preventing disengagement due to relapse by providing features that meet users' need for stimulation should therefore be explored.

Fourthly, although findings from focus groups with young adults who drink at hazardous or harmful levels indicate a strong preference for features that foster relatedness [326], evidence from studies with adult drinkers suggests that people typically react differently to features that connect them with friends or other users [287]. These results suggest that excessive drinkers may either strongly like or dislike social support features or challenge features.

6.5.2 Limitations

This study was limited by employing an abstract, cognitively demanding ranking task that may have been more suitable for a face-to-face (as opposed to an online) study context. A plausible explanation as to why goal setting to use the app was ranked highly in the online sample is that users thought that this referred to goal setting for alcohol reduction. Potential misunderstandings were mitigated by careful piloting of the feature descriptions, but it is possible that some participants were still confused. Although participants' rankings should be interpreted with caution, the qualitative findings aid in the interpretation of the quantitative results.

It has been argued that users find it difficult to discuss design concepts without visual or tactile prompts, or that users are not designers [327]. Indeed, some participants in the present study found it difficult to articulate concrete design suggestions, such as how a narrative linked to alcohol reduction would pan out. However, an abstract ranking task was deemed most suitable to avoid limiting participants' imagination of particular features.

It is possible that the labels used for the engagement features may have biased participants' attitudes. This is suggested by a study in which old adults agreed that a 'falls-prevention intervention' was a good idea, but only for people who were older or frailer than them. The authors of the study therefore concluded that reframing the intervention as a 'balance-training programme' might promote uptake [328]. In the present study, labels such as 'professional support features' may have been perceived as too serious or irrelevant to participants' particular situations. This was suggested by a few participants in the focus groups. It is therefore possible that the finding that professional support features were preferred by participants who identified as being a 'heavy' drinker is an artefact of the labels used.

As men tend to exhibit more alcohol-related problems than women across countries [329,330], the recruitment of more women than men into the focus groups constitutes a limitation. Future research should attempt to recruit a more balanced sample, with a view to exploring possible gender differences in app preferences. However, it should be noted that just over half of the online sample were male, and that no differential preferences based on gender were identified in this sample. Moreover, while the current approach to eliciting user needs provides useful information, an experimental study, in which the presence or design of particular features is manipulated, is required in order to test the actual impact on app engagement.

6.5.3 Conclusion

There was low agreement between participants concerning the importance of particular engagement features, but on average, those judged to be most important for inclusion in smartphone apps for alcohol reduction were personalisation, interactive features and control features. This study highlights that different features may be liked and used by different users, which should be considered in the design of novel alcohol reduction apps, or the modification of existing ones.

6.5.4 Citation for the published peer-reviewed article for this study

Perski, O., Baretta, D., Blandford, A., West, R., & Michie, S. (2018). Engagement features judged by excessive drinkers as most important to include in smartphone applications for alcohol reduction: A mixed-methods study. *Digital Health, 4,* 1-15. https://doi.org/10.1177/2055207618785841

See Appendix 15 for the published peer-reviewed journal article.

6.5.5 Next steps

Findings from the studies reported in Chapters 2, 3 and 5 suggested that motivation to change, perceived usefulness of the app, the target behaviour itself and perceived lack of time may influence engagement with apps for alcohol reduction. However, these studies were limited by relying on participants' ability to predict their future preferences, experiences and behaviour. It was therefore considered important to triangulate these findings with behavioural data. Study 6 (reported in Chapter 7) employed a series *N*-of-1 designs to examine within-subjects (as opposed to between-subjects) predictors of engagement with the *Drink Less* app, harnessing real-time Ecological Momentary Assessments.

7 CHAPTER 7 – Do daily fluctuations in psychological and app-related variables predict within-person variability in engagement with an alcohol reduction app? A series of *N*-of-1 designs (Study 6)

7.1 Abstract

Background: Previous studies have highlighted between-subjects predictors of engagement with apps for alcohol reduction, including psychological (e.g. motivation to change) and app-related (e.g. the receipt of a daily reminder) variables. However, strategies to promote engagement need to be effective at the individual level. Evidence as to whether these between-subjects predictors of engagement are also predictive for individuals is lacking.

Purpose: To examine whether daily fluctuations in i) the receipt of a daily reminder, ii) motivation to reduce alcohol, iii) perceived usefulness of the app, iv) alcohol consumption, and v) perceived lack of time are predictive of within-person variability in the frequency and amount of engagement with a theory- and evidence-based alcohol reduction app, *Drink Less.*

Methods: This study used a series of observational *N*-of-1 designs. Psychological and app-related predictor variables were measured using twice-daily Ecological Momentary Assessments for 28 days, sent to participants via text messages. The outcome variables (i.e. the frequency and amount of engagement) were measured objectively through automated recordings of participants' app screen views. Nine London-based adults who drank alcohol excessively and were willing to set a goal to reduce their drinking took part.

Each participant's dataset was analysed separately using Generalised Additive Mixed Models.

Results: Different variables were significant predictors of the frequency and amount of engagement within and between individuals. The receipt of a daily reminder (IRRs = 1.80-3.88, *p*'s < .05) and perceived usefulness of the app (IRRs = 0.82-1.42, *p*'s < .05) were the most consistent predictors of within-person variability in the frequency of engagement. Motivation to reduce alcohol (IRRs = 1.67-3.45, *p*'s < .05) and perceived usefulness of the app (IRRs = 0.52-137.32, *p*'s < .05) were the most consistent predictors of within-person variability in the amount of engagement.

Conclusion: The utility of the app-related and psychological variables in predicting the frequency and amount of engagement with the *Drink Less* app differed within and between individuals. This suggests that different strategies to promote engagement may be required for different individuals, and that such strategies may have differential effects on the different facets of engagement.

7.2 Introduction

Studies to date have typically focused on the identification of between-subjects predictors of engagement with DBCIs for alcohol reduction [281]. As strategies to increase engagement also need to be effective for individuals [331,332], it is important to examine whether key predictors identified at the between-subjects level are also predictive of engagement at the individual level. This study aimed to assess within-person predictors of the frequency (i.e. number of logins) and amount (i.e. time spent per login) of engagement with a theory- and evidence-based alcohol reduction app, *Drink Less* [95,97].

Published secondary analyses of data from RCTs of web- and app-based interventions for alcohol reduction have typically been used to identify between-subjects predictors of engagement. These studies show that demographic characteristics, such as being female, older and more highly educated, are positively associated with the frequency and amount of engagement [164,333,334]. Higher baseline levels of motivation to change [164,335] and lower baseline levels of alcohol consumption [164,310,333] have been found to predict the frequency of engagement. Moreover, app-related variables, such as the receipt of proactive reminders, have also been found to promote the frequency of engagement [28]. Studies 2 and 5 reported in this thesis (Chapters 3 and 6) prompted excessive drinkers to reflect on factors they expected to be most important for engagement with apps for alcohol reduction. These studies highlighted that motivation to change, perceived personal relevance of the app (defined as the extent to which the user believes that the app is suited to their individual needs [287]), and perceived usefulness of the app (defined as the extent to which the individual believes that use of the app will enhance task performance [52]), were judged to be important for engagement [287,336]. Although common themes were identified in these studies, agreement between potential users on what factors were expected to be most important for engagement was low [336]. Qualitative research has also been conducted with participants who disengaged prior to completion of an RCT of a web-based alcohol reduction intervention [164]. When asked to retrospectively report on why they disengaged from the intervention, users frequently mentioned perceived lack of time (e.g. being too busy, having other priorities), dissatisfaction with the intervention (e.g. poor usability, irrelevant content) and improvement in the condition (e.g. feeling better).

Quantitative studies examining predictors of engagement have typically relied on between-subjects designs, aggregating data across participants. However, individual-level interventions, including apps for alcohol reduction, are designed to target within-subjects processes that lead to behaviour change. For intervention strategies aimed at increasing engagement (e.g. proactive reminders, rewards, feedback) to be promoted, they need to be shown to be effective at the individual, not just at the group level. It is therefore important to assess whether associations identified at the between-subjects level are also identified at the within-subjects level. The *N*-of-1 design, also known as a single-case design, is ideally suited for the assessment of within-person processes. The *N*-of-1 design can be either observational or experimental and "…receives its name by virtue of its sample size: *N* is equal to one" [337].

Qualitative studies have relied on either prospective or retrospective (as opposed to real-time) self-reports of psychological processes. Such reports are likely to be biased or inaccurate [338]. For example, when prospectively predicting what factors are expected to be most important for engagement, potential users tend to highlight app-related aspects, such as the presence of features that enhance motivation to change (e.g. goal setting, self-monitoring and proactive reminders) and perceived usefulness (e.g. tailoring of content, rewards) [287,336]. However, when asked to retrospectively report on what factors participants think contributed to their disengagement from an intervention, different aspects tend to be highlighted, such as perceived lack of time [164]. Ecological Momentary Assessment (EMA) is a data gathering method which overcomes the problems associated with both prospective and retrospective self-reports, as it allows the examination of psychological processes simultaneously to the behaviour (i.e. in 'real-time') [226,339].

The present study used a series of *N*-of-1 designs, harnessing EMAs, to examine whether daily fluctuations in i) the receipt of a reminder, ii) motivation to reduce alcohol, iii) perceived usefulness of the app, iv) alcohol consumption, and v) perceived lack of time are predictive of within-person variability in the frequency and amount of engagement with a theory- and evidence-based alcohol reduction app, *Drink Less*. The decision to focus on 'perceived usefulness of the app' in the present study (as opposed to 'perceived relevance of the app') was informed by a meta-analysis of 59 studies indicating that the variable 'perceived usefulness' is consistently associated with behavioural intentions to use technology (r = 0.59) [53]; less is known about the relationship between the variable 'perceived relevance' and key outcome variables. By measuring predictor variables prior to the measurement of the outcome variables, this study aimed to provide a greater understanding of the temporal direction of the relationships under investigation.

7.3 Methods

7.3.1 Study design

A pre-registered study protocol can be found on the OSF (osf.io/zn79m). This study used a series of observational *N*-of-1 designs with twice-daily (i.e. morning and evening) self-report measures of cognitive and behavioural predictor variables. The dependent variables were the objectively recorded frequency and amount of engagement with the *Drink Less* app, described in detail below. Although the subjective experience during app use (i.e. attention, enjoyment, interest) is also thought to be necessary for someone to be engaged (see Chapters 2 and 4), it was considered important to minimise participant burden. Hence, only behavioural indicators of engagement were considered, as these could be measured automatically via participants' app screen views. The key outcome of interest, specified in the pre-registered analysis plan, was the 'frequency of engagement' (i.e. the number of logins per measurement period). To complement these analyses, the variable 'amount of engagement' (i.e. the time spent on the app per measurement period) was also investigated in a series of unplanned analyses.

7.3.1.1 Participants

7.3.1.1.1 Eligibility criteria

Participants were eligible to take part if they i) were aged 18+ years; ii) owned an iPhone capable of running iOS v.8.0 software or higher (i.e. iPhone 4S or later models); iii) resided in or near London (UK) and were willing to come into University College London (UCL) on one occasion (to ensure adequate study commitment); iv) reported an AUDIT score of ≥ 8, indicating excessive alcohol consumption [77]; v) were interested in using an app to reduce their drinking; vi) were willing to set a goal to reduce their drinking; vii) installed the *Drink Less* app and opened it at least once following the briefing interview (see 'Procedure' section below for more details); viii) were willing to engage with the app daily for 28 days, recognising that there may be occasional days where they would not engage with it [340]; and ix) were willing to respond to twice-daily text messages for 28 days. Participants were excluded if they were not fluent English speakers.

7.3.1.1.2 Sampling

Participants were recruited online through Call for Participants (www.callforparticipants.com), social media (e.g. Twitter) and charitable alcohol reduction organisations' mailing lists. The recruitment materials stated that drinkers were invited to take part in a study about how people use apps for alcohol reduction in their daily lives, which involved responding to twice-daily text messages for 28 days (see Appendix 12).

7.3.1.1.3 Sample size

The number of observations (and not the number of participants) determine statistical power in *N*-of-1 designs [341]. As the *Drink Less* app was designed to be used for 28 days, each participant was asked to respond to twice-daily EMAs for 28 days (i.e. up to 56 observations per participant). The measurement frequency of two EMAs per day was informed by prior research conducted within the behavioural science domain [342]. As data were planned to be analysed using Generalised Additive Mixed Models (see the 'Data analysis' section below for more details), Monte Carlo simulations [343] were run to estimate the statistical power that would be achieved with a total of 56 data inputs. A simulation-based power analysis conducted in R indicated that this study would have 80% power to detect an incident rate ratio (IRR) of 1.8 for the association between 'perceived usefulness of the app' (predictor variable) and 'frequency of engagement' (outcome variable). Given uncertainties regarding the distribution of model parameters, this power analysis should be interpreted with caution. Details about statistical assumptions used to inform the power analysis are shown in Table 7.1. To allow qualitative (but not quantitative) assessment of potential between-subjects differences in the associations between the predictor variables and app engagement, a total of 8 participants was considered to be sufficient [342,344,345]. As previous *N*-of-1 studies report up to 47% study drop-out [342,344,345], the researcher aimed to recruit an additional 50% of the target sample (i.e. 12 participants).

Table 7.1. Statistical assumptions used to inform the simulation-based power analysis

Considerations	Statistical assumptions and source of information (where available)
Model type	Generalised Additive Mixed Model (GAMM)
Number of observations	Twice-daily EMAs for a period of 28 days (i.e. a total of 56 observations per participant).
Seasonality	No seasonality reflected by the day of the week the data were collected.
Distribution and point estimate (dependent variable)	The dependent variable (i.e. 'frequency of engagement', operationalised as the number of app logins per measurement period) was assumed to follow a Poisson distribution with a mean of 11.7 logins per measurement period [95]. As the dependent variable represents count data, it was expected to follow a Poisson distribution. The mean of 11.7 logins was drawn from a between-subjects, factorial RCT of the *Drink Less* app [95], as this was judged to represent the best available data.
Distribution and point estimate (independent variable)	The independent variable (i.e. 'perceived usefulness of the app') was assumed to follow an Auto-Regressive (AR) Integrated Moving Average process with first-order autocorrelation, as it was expected that measurements would be similar to those taken 12 hours previously. As prior studies in the alcohol reduction domain assessing the variable 'perceived usefulness of the app' are lacking, we drew on results from the between-subjects, factorial RCT of the *Drink Less* app, which assessed the variable 'helpfulness of the app' at 28-day follow-up. This variable was deemed to be conceptually similar to the target variable. It was therefore assumed that the mean level of the independent variable would be 3.18 (*SD* = 0.93) [95].

7.3.1.2 <u>Measures</u>

7.3.1.2.1 Online screening questionnaire

The following data were collected at baseline to determine study eligibility and to describe the sample (see Appendix 13): i) age; ii) gender; iii) type of work (i.e. manual, non-manual, other); iv) whether participants owned an iPhone capable of running iOS 8.0 software or higher (i.e. iPhone 4S or later models); v) whether participants were residing in or near London and were willing to come into UCL for a briefing interview (yes vs. no); vi) alcohol consumption, measured using the AUDIT; vii) whether participants were interested in using an app to reduce their drinking (yes vs. no); viii) whether participants were willing to set a goal to reduce their drinking (yes vs. no); ix) whether participants were willing to engage with the study app daily for 28 days (yes vs. no); x) whether participants had previously used an alcohol reduction app (yes vs. no) and if so, which one; and xi) whether participants were willing to respond to the twice-daily text messages for 28 days (yes vs. no).

7.3.1.2.2 Ecological Momentary Assessments (predictor variables)

The following data were collected twice per day (i.e. morning and evening) via text messages, sent manually from an iPhone 6S by the researcher (see Appendix 14):

1. ‘Motivation to reduce alcohol’, measured by asking: “How motivated are you currently to reduce your drinking?”. The response options ranged from 1-7, with 1 indicating ‘not at all’ and 7 indicating ‘extremely’.

2. ‘Perceived usefulness of the app’, measured by asking: “How useful do you currently think the *Drink Less* app is for you?” The response options ranged from 1-7, with 1 indicating ‘not at all’ and 7 indicating ‘extremely’.

3. ‘Alcohol consumption’, measured by asking: “How many drinks containing alcohol have you had in the past 12 hours?” Participants were asked to input an integer from 0 (no drinks) to infinity (i.e. whole drinks). To minimise respondent burden and because absolute amount of alcohol was not of interest in the present study, participants were not asked to consider the number of units consumed. Hence, participants were asked to enter ‘1’ even if they had only had half a pint of beer.

4. ‘Perceived lack of time’, measured by asking participants: “To what extent do you currently have time for the *Drink Less* app?” The response options ranged from 1-7, with 1 indicating ‘I don’t have any time for the app’ and 7 indicating ‘I have lots of time for the app’.

7.3.1.2.3 Additional predictor variables (tailored to participants’ preferences)

5. Whether or not a proactive reminder was received during each 12-hour period. This variable was coded 1 if a reminder was received and 0 if it was not received. The maximum number of reminders received every 24 hours was 1 (limited by the core design of the study app), and this pattern did not change during the course of the study. This variable was tailored to participants’ preferences at the outset of the study (i.e. whether or not participants wanted to have the reminder switched on or

off throughout the study). More details about the reminder are reported in the 'Procedure' section below.

7.3.1.2.4 Outcome variables

App screen views were automatically recorded, stored in an online database and extracted using the free python library *pandas* (https://pandas.pydata.org/) to calculate objective frequency and amount of engagement. The variable 'frequency of engagement' was operationalised as the number of logins during each 12-hour measurement period (e.g. 10AM to 10PM; 10PM to 10AM), with a login defined as a new screen view following at least 30 minutes of inactivity [302]. The variable 'amount of engagement' was derived by calculating the time spent (in seconds) per measurement period. For descriptive purposes, the variable 'depth of engagement' was also derived, which was operationalised as the number of app components accessed per measurement period, indexed as a proportion of the number of available components (i.e. Goal Setting; Self-monitoring/Feedback; Action Planning; Normative Feedback; Cognitive Bias Re-Training; Identity Change; Other [95]). However, as 'depth of engagement' was strongly correlated with 'amount of engagement' for all participants, no inferential analyses were conducted using this variable.

7.3.1.3 Intervention

The *Drink Less* app is a stand-alone DBCI designed to promote alcohol reduction in adults who drink excessively. The app is centred around a goal setting module which allows users to select one or multiple weekly goals of their choice (e.g. maximum number of units, alcohol-free days, maximum spending on alcohol, or maximum number of alcohol-attributed calories). The app also includes five additional intervention modules: i) Normative Feedback, ii) Cognitive Bias Re-Training, iii) Self-Monitoring and Feedback, iv) Action Planning, and v) Identity Change. Details about how intervention content was selected [86,98], user feedback on a first version of the app [346], the development process [97] and a first evaluation of the app's components in a factorial RCT [95] have been described in detail elsewhere. The *Drink Less* app allows users to set one daily reminder to open the app, which can be switched on or off depending on the user's preferences, and set to a suitable timing.

7.3.1.4 Procedure

Participants who expressed an interest in taking part were asked to read the participant information sheet, provide informed consent and fill out the online screening questionnaire.

Eligible participants were invited to a briefing interview at UCL where they were asked to re-read the information sheet and were consented. Participants were asked to download the *Drink Less* app, briefly explore it, and set at least one weekly alcohol reduction goal of their choice. Participants were asked if they wanted to switch the daily reminder on or off and if applicable, select a suitable timing for these. After having explored the app, participants were asked to complete a brief survey on their phone, which fetched their unique user ID, generated by the *Drink Less* app. This information enabled the researchers to match participants to their app screen views and hence, derive the outcome variables. Participants were asked a few questions about their expected app use and what they were hoping to achieve using the app. Participants were subsequently asked to select suitable timings for the twice-daily text messages. In the morning, participants were asked to select a suitable time between 6am and 10am; and in the evening, between 6pm and 10pm, ensuring that the selected time points did not fall earlier/later than their usual morning and evening bedtimes, respectively. Participants were subsequently asked to familiarise themselves with the daily EMA questions and response options, and practised inputting their responses to the four questions into a single text message. No particular instructions about app engagement were provided other than that participants were expected to engage with the app at least once daily for 28 days, recognising that there may be occasional days when they would not engage with it. Participants were told that they had to respond to at least 70% of the text messages and take part in the debriefing interview to receive any payment. They were also asked to notify the researcher if they decided to change the timing of the daily reminder, so that this could be accounted for in the statistical analyses. Participants were told that they could not change the timing of the twice-daily text messages over the course of the study, so as not to complicate the statistical analyses. The briefing interviews lasted between 29 to 63 minutes.

After the briefing interview, participants were asked to respond to the twice-daily text messages for 28 days, sent to users via a single text message. The first text message was sent the morning after the briefing interview. When a response was received, participants were sent the following standard response: "Thank you for your responses!". Participants also received weekly updates via text message about their survey response rates to encourage adherence to the study materials (e.g. "Hi X! Thank you for completing the first week of the study. You have responded to X/14 text messages. Keep up the good work!"). If the text messages were not received in the expected format, participants received a standard reply with instructions for how to input the responses (see Appendix 14).

After 28 days, participants were invited to take part in a debriefing interview conducted over the phone, during which they were asked about the acceptability of the twice-daily text messages and their experiences of engaging with the app. The debriefing interviews lasted between 25 to 47 minutes. Participants were paid £0.5 per data input (i.e. a maximum of £28), in addition to £32 upon study completion (i.e. £60 in total). This was paid to participants in the form of a shopping voucher.

*7.3.1.5 **Data analysis***

Guided by published research in the behavioural science domain [342,344,345], in time series with > 5% missing data, multiple imputation was carried out using an expectation-maximisation with bootstrapping algorithm via the R package *Amelia II*. Data were imputed separately for each dataset (i.e. each participant). A polynomial time trend (e.g. linear, quadratic) was included if this was found to improve the precision of the imputed data points. This was decided upon by examining the 95% confidence intervals (CIs) of the means of the imputed data points. Five imputed datasets were created per dataset with missing values, which were combined prior to further statistical analyses using Rubin's rules [342,344,345].

Descriptive statistics were calculated for each participant. Time series analyses were conducted via the R package *mgcv*: Generalised Additive Mixed Models

(GAMMs) were fitted to estimate incident rate ratios (IRRs) for the predictor variables. The IRR is a measure of relative difference and can, in this particular context, be interpreted as the relative rate of logins or amount of engagement for the different levels of the predictor variables. The GAMM is a type of multilevel model which has previously been applied to data from *N*-of-1 designs. GAMMs are particularly well-suited to the modelling of time series data with one level of measurement (i.e. repeated measurements nested within one individual), as it can accommodate the inclusion of autocorrelated error terms [347]. The analyses proceeded in a number of stages using a 'backwards' selection procedure:

1. As the outcome variables represented counts, data were first assessed for overdispersion (i.e. when the variance is greater than the mean). If there was evidence for overdispersion, a quasi-Poisson distribution (as opposed to a Poisson distribution) was specified.

2. As repeated measures taken from the same individual are often correlated, data from *N*-of-1 studies typically violate the assumption of independence of observations. Autocorrelation was therefore assessed through the autocorrelation function and the partial autocorrelation function. For example, evidence of first-order autocorrelation means that measurements are significantly correlated with those taken 12 hours previously.

3. Parsimonious models were subsequently built for each participant through the stepwise elimination of redundant terms: a full model including all predictor variables was first fitted to determine the most

appropriate autocorrelation structure for each participant. Model fit was compared using Akaike's Information Criterion [348]. The predictor variables were then sequentially varied to arrive at a best fitting model for each participant. Although the *a priori* power analysis did not take account of adjustment for seasonality or moving average terms, it was determined *a posteriori* that adjustment for the day of the week through the inclusion of a cyclic cubic smoothing term significantly improved the model fit for all participants and that the inclusion of a moving average term improved the model fit for some participants.

7.3.1.6 Ethical approval

Ethical approval was granted by UCL's Computer Science Departmental Research Ethics Chair (Project ID: UCLIC/1617/004/Staff Blandford HFDH). Personal identifiers were removed, and anonymised data were stored securely on a password protected computer. Participants' contact details were stored separately in a locked cabinet. The SIM card used to deliver the daily text messages was wiped when data collection had been completed.

7.4 Results

7.4.1 Participants

Of 22 participants who completed the online screening questionnaire, 11 met the inclusion criteria and were invited to take part. One participant was unable to initiate the 28-day study during the allocated study period. In total, ten participants took part in the study between June 29th and August 9th 2018. One participant broke their phone 14 days into the study and re-downloaded the app onto a new phone without notifying the researcher. Due to technical issues, the new phone's app screens failed to sync with the database and hence, the outcome data for the last 14 days of the study were lost. Hence, this participant was excluded from the inferential analyses. In line with the desired sample size of 8 participants, a total of 9 participants were included in the inferential analyses (although descriptive statistics were calculated for all 10 participants).

Participants' characteristics are summarised in Table 7.2. The mean age of participants was 25.0 years (*SD* = 4.4), 90% were female and the majority were employed in a non-manual profession (60%). Participants had a mean AUDIT score of 15.5 (*SD* = 7.2), indicating excessive alcohol consumption. Two participants had prior experience of using an alcohol reduction app, with one participant having prior experience of using the *Drink Less* app. As participants serve as their own comparisons in *N*-of-1 designs, the participant who had previously used the *Drink Less* app was still invited to take part in the study.

Table 7.2. Participants' demographic, drinking and app-related characteristics.

ID	**Gender**	**Age**	**Type of work**	**AUDIT**	**Prior use of an alcohol reduction app**	**Prior use of the *Drink Less* app**
P1	Female	28	Non-manual	16	No	No
P2	Female	20	Other	10	No	No
P3	Female	25	Non-manual	30	No	No
P4	Female	18	Other	12	No	No
P5	Male	21	Other	22	No	No
P6	Female	31	Non-manual	8	No	No
P7	Female	23	Non-manual	12	Yes	Yes
P8	Female	30	Non-manual	11	No	No
P9	Female	28	Other	23	Yes	No
P10	Female	26	Non-manual	10	No	No

7.4.2 Descriptive statistics

Eight participants (80%) decided to have the daily reminder switched on during the study, with two participants (20%) deciding to have it switched off. Overall, participants displayed high adherence to the daily text messages (*M* = 93%, *SD* = 5.8%), with the number of missing responses varying from 0-16%. The mean level of the psychological predictor variables over the course of the study varied across participants (see Table 7.3).

Table 7.3. Adherence to the twice-daily text messages and descriptive statistics for the predictor variables.

ID	**Adherence, *N* (%)**	**Timing of text messages**	**Daily reminder switched on/off**	**Timing of daily reminder**	**Motivation to reduce alcohol, M (*SD*); range**	**Perceived usefulness of the app, M (*SD*); range**	**Alcohol consumption, M (*SD*); range**	**Perceived lack of time, M (*SD*); range**
P1	56 (100%)	10AM/PM	ON	10AM	5.3 (*1.1*); 3-7	5.4 (*0.8*); 4-7	2.1 (*2.8*); 0-10	6.1 (*1.2*); 3-7
P2	55 (98%)	10AM/PM	ON	1PM	6.3 (*1.1*); 3-7*	6.3 (*1.1*); 3-7*	0.1 (*0.5*); 0-3*	4.6 (*2.2*); 1-7*
P3	50 (89%)	7.30AM/PM	ON	4PM	5.2 (*0.9*); 4-7*	5.3 (*1.1*); 3-7*	1.2 (*1.3*); 0-5*	4.5 (*1.0*); 2-7*
P4	49 (87.5%)	10AM/PM	ON	11AM	4.1 (*1.6*); 1-7*	2.4 (*1.3*); 1-5*	0.1 (*0.8*); 0-4*	4.9 (*1.8*); 2-7*
P5	55 (98%)	9.30AM/PM	OFF	-	3.6 (*1.0*); 2-6*	3.6 (*1.2*); 1-7*	1.2 (*1.7*); 0-8*	3.9 (*0.9*); 2-7*
P6	47 (84%)	10AM/PM	ON	10AM	5.6 (*0.7*); 4-7*	4.4 (*0.6*); 4-6*	0.3 (*0.8*); 0-3*	4.4 (*0.7*); 3-7*
P7	48 (86%)	9AM/PM	ON	9AM	4.1 (*1.2*); 1-6*	3.2 (*0.9*); 2-5*	1.1 (*2.1*); 0-6*	2.8 (*1.6*); 1-6*
P8	51 (91%)	10AM/PM	OFF	-	5.9 (*0.5*); 4-7*	6.1 (*0.9*); 4-7*	0.4 (*0.9*); 0-4*	2.2 (*1.4*); 1-5*
P9	56 (100%)	10AM/PM	ON	10.30AM	4.3 (*1.9*); 1-7	1.9 (*0.9*); 1-5	3.9 (*4.3*); 0-14	6.0 (*1.3*); 2-7
P10	54 (96%)	10AM/PM	ON	9AM	5.3 (*1.6*); 1-7*	4.8 (*1.0*); 1-6*	1.9 (*2.9*); 0-9*	5.5 (*1.0*); 3-7*

Note. * For participants with missing data, means and standard deviations for the complete datasets (after multiple imputation) were computed using Rubin's rules.

Over the course of the study, participants' total number of logins ranged from 10-69 (see Figure 7.1 for plots of participants' frequency of engagement over the course of the study). The total depth of use ranged from 0.14 (i.e. accessing one of the app's seven components) to 0.86 (i.e. accessing six of the app's seven components). The total amount of use ranged from 4 minutes and 24 seconds to 70 minutes and 14 seconds. The average amount of use per measurement period ranged from 0 minutes and 0 seconds (i.e. measurement periods with no logins) to 1 minute and 15 seconds (see Table 7.4).

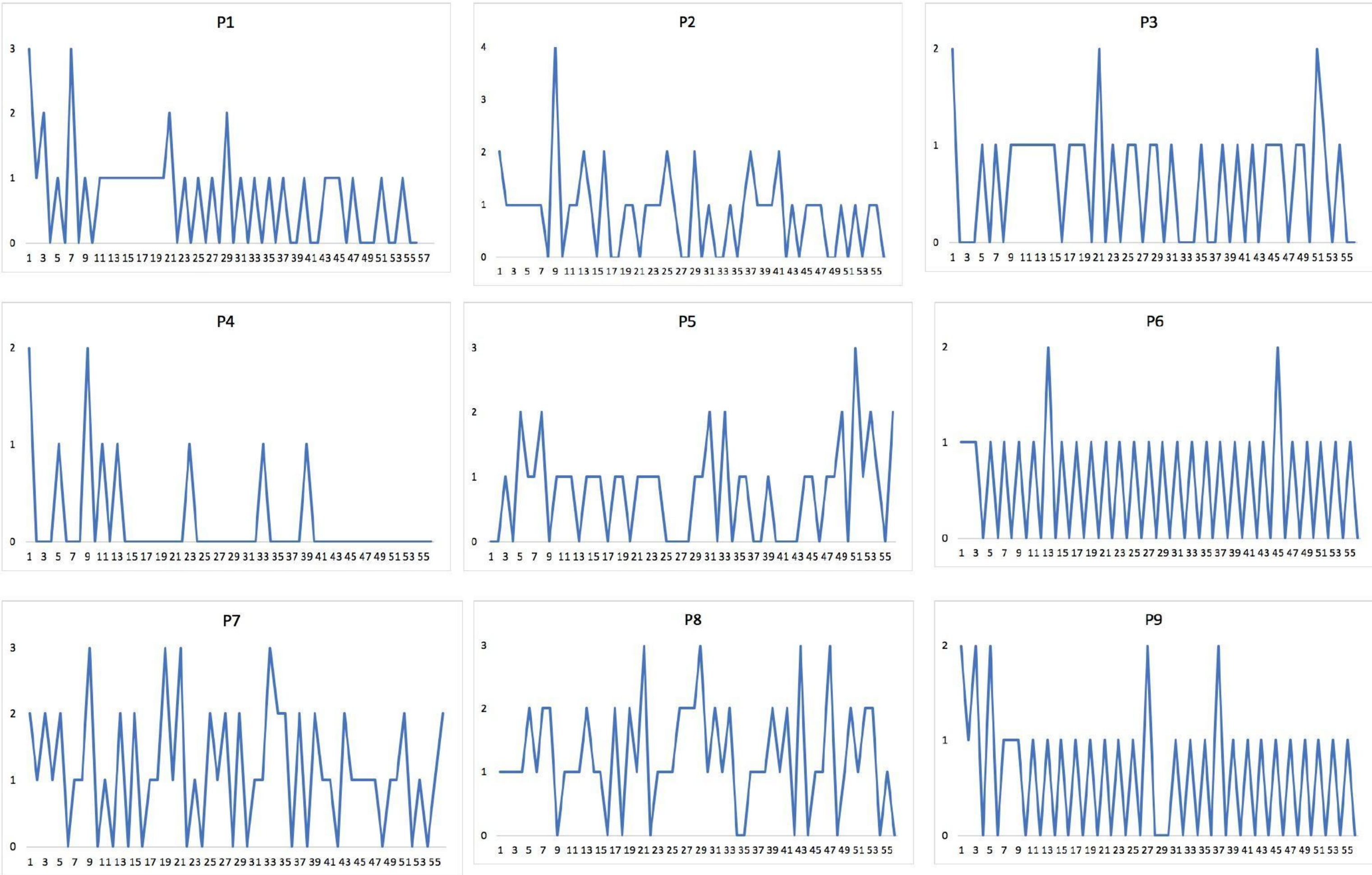


Figure 7.1. Plots of participants' frequency of engagement over the course of the study period.

Table 7.4. Descriptive statistics of participants' frequency, amount and depth of engagement with the *Drink Less* app.

ID	Total number of logins; M (*SD*); range	Total depth of engagement	Depth of engagement per login, M (*SD*)	Total amount of engagement (mm:ss)	Amount of engagement per login (mm:ss), M (*SD*); range
P1	39; 0.7 (*0.7*); 0 – 3	0.71	0.10 (*0.12*)	23:11	00:25 (*00:48*); 00:00 – 02:49
P2	47; 0.8 (*0.8*); 0 – 4	0.86	0.20 (*0.20*)	60:43	01:05 (*02:32*); 00:00 – 16:32
P3	35; 0.6 (*0.6*); 0 – 2	0.57	0.10 (*0.11*)	13:12	00:14 (*00:27*); 00:00 – 02:19
P4	10; 0.2 (*0.5*); 0 – 2	0.43	0.03 (*0.08*)	04:24	00:05 (*00:18*); 00:00 – 01:29
P5	42; 0.8 (*0.7*); 0 – 3	0.29	0.11 (*0.11*)	18:20	00:20 (*00:29*); 00:00 – 02:11
P6	31; 0.6 (*0.6*); 0 – 2	0.57	0.09 (*0.11*)	39:19	00:42 (*01:25)*; 00:00 – 08:12
P7	64; 1.1 (*0.9*); 0 – 3	0.14	0.10 (*0.06*)	19:14	00:21 (*00:25)*; 00:00 – 02:24
P8	69; 1.2 (*0.9*); 0 – 3	0.43	0.17 (*0.13*)	70:14	01:15 (*02:08*); 00:00 – 10:47
P9	34; 0.6 (*0.7*); 0 – 2	0.43	0.09 (*0.11*)	35:26	00:38 (*02:04*); 00:00 – 13:40
P10	N/A*	N/A*	N/A*	N/A*	N/A*

Note. * Due to a technical issue, outcome data were lost for one participant.

7.4.3 Predicting the frequency and amount of engagement

Table 7.5 reports the Generalised Additive Mixed Models in which the frequency and amount of engagement were each regressed onto the predictor variables, adjusting for the day of the week, autocorrelation and moving average terms. For two participants (P3 and P8), none of the predictor variables assessed was significantly associated with the number of logins. None of the independent variables assessed was significantly associated with the amount of engagement for two participants (P1 and P8).

7.4.3.1 Daily reminder

The receipt of a reminder was a significant predictor of the number of logins for three participants (IRRs = 1.80-3.88, all *p*'s < .05). For these participants (P1, P6 and P7), the receipt of a reminder was associated with an 80-288% increase in the number of logins in the next 12 hours.

The receipt of a reminder was a significant predictor of the amount of engagement for one participant (IRR = 4.31, 95% CI = 1.73-10.73, *p* < .01). For this participant (P3), the receipt of a reminder was associated with a 331% increase in the amount of engagement in the next 12 hours.

7.4.3.2 Motivation to reduce alcohol

Motivation to reduce alcohol was a significant predictor of the number of logins for one participant (IRR = 1.14, 95% CI = 1.02-1.27, *p* = .02). For this participant (P4), a 1-point increase in motivation to reduce alcohol was associated with a 14% increase in the number of logins in the next 12 hours.

Motivation to reduce alcohol was a significant predictor of the amount of engagement for three participants (IRRs = 1.67-3.45, all *p*'s < .05). For these participants (P4, P6 and P7), a 1-point increase in motivation to reduce alcohol was associated with a 67-245% increase in the amount of engagement in the next 12 hours.

7.4.3.3 Perceived usefulness of the app

Perceived usefulness of the app was a significant predictor of the number of logins for three participants (IRRs = 0.82-1.42, all *p*'s < .05). For one participant (P1), a 1-point increase in perceived usefulness of the app was associated with an 18% reduction in the number of logins in the next 12 hours. For two participants (P5 and P9), a 1-point increase in perceived usefulness of the app was associated with a 38-42% increase in the number of logins in the next 12 hours.

Perceived usefulness of the app was a significant predictor of the amount of engagement for four participants (IRRs = 0.52-137.32, all *p*'s < .05). For one participant (P7), a 1-point increase in perceived usefulness of the app was associated with a 48% reduction in the amount of engagement in the next 12 hours. For three participants (P4, P5 and P9), a 1-point increase in perceived usefulness of the app was associated with a 67-13,632% increase in the amount of engagement in the next 12 hours.

7.4.3.4 Alcohol consumption

The number of alcoholic drinks consumed in the past 12 hours was a significant predictor of the number of logins for one participant (IRR = 1.50, 95% CI = 1.16-

1.93, $p < .01$). For this participant (P2), each alcoholic drink consumed in the past 12 hours was associated with a 50% increase in the number of logins in the next 12 hours.

The number of alcoholic drinks consumed in the past 12 hours was a significant predictor of the amount of engagement for two participants (IRRs = 1.38-2.38, p's < .01). For these participants (P2, P3), each alcoholic drink consumed in the past 12 hours was associated with a 38-138% increase in the amount of engagement in the next 12 hours.

7.4.3.5 Perceived lack of time

Perceived lack of time was a significant predictor of the number of logins for two participants (IRRs = 0.77-1.13, p's < .05). For one participant (P6), a 1-point increase in perceived lack of time (meaning that they had more time for the app) was associated with a 23% reduction in the number of logins in the next 12 hours. For the other participant (P2), a 1-point increase in perceived lack of time was associated with a 13% increase in the number of logins in the next 12 hours.

Perceived lack of time was a significant predictor of the amount of engagement for two participants (IRRs = 0.20-4.77, p's < .05). For one participant (P4), a 1-point increase in perceived lack of time (meaning that they had more time for the app) was associated with an 80% reduction in the amount of engagement in the next 12 hours. For the other participant (P9), a 1-point increase in perceived lack of time was associated with a 377% increase in the amount of engagement in the next 12 hours.

Table 7.5. Incident rate ratios (IRRs) for the associations between the variability in the predictor variables and variability in the frequency and amount of engagement for each participant.

	Frequency of engagement		**Amount of engagement**	
	IRR (95% CI)	***p*-value**	**IRR (95% CI)**	***p*-value**
P1				
Reminder	**$1.80_{2,1}$ (1.19-2.74)**	**.01***	-	-
Motivation to reduce alcohol	**$1.14_{2,1}$ (1.02-1.27)**	**.02***	$1.12_{0,0}$ (0.68-1.83)	.65
Perceived usefulness of the app	**$0.82_{2,1}$ (0.68-0.99)**	**.04***	-	-
Alcohol consumption	-	-	-	-
Perceived lack of time	$0.93_{2,1}$ (0.86-1.02)	.15	-	-
P2				
Reminder	$1.99_{1,0}$ (0.67-5.94)	.22	-	-
Motivation to reduce alcohol	-	-	-	-
Perceived usefulness of the app	-	-	-	-
Alcohol consumption	**$1.50_{1,0}$ (1.16-1.93)**	**< .01****	**$2.38_{1,0}$ (1.65-3.43)**	**< .01****
Perceived lack of time	**$1.13_{1,0}$ (1.01-1.25)**	**.03***	-	-
P3				
Reminder	-	-	**$4.31_{0,0}$ (1.73-10.73)**	**< .01****
Motivation to reduce alcohol	$0.89_{1,0}$ (0.67-1.19)	.45	-	-
Perceived usefulness of the app	-	-	-	-
Alcohol consumption	-	-	**$1.38_{0,0}$ (1.11-1.73)**	**< .01****
Perceived lack of time	-	-	$1.19_{0,0}$ (0.79-1.77)	.40

Note. All models were adjusted for the day of the week using a cyclic cubic smoothing term. Numbers in subscript indicate the lags of autoregressive (AR) and moving average (MA) terms, respectively. A lag value of 0 indicates that an AR or MA term was not included; * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 7.5. *Continued.*

	Frequency of engagement		Amount of engagement	
	IRR (95% CI)	***p*-value**	**IRR (95% CI)**	***p*-value**
P4				
Reminder	-	-	-	-
Motivation to reduce alcohol	**$1.88_{0,0}$ (1.22-2.91)**	**< .01****	**$2.03_{0,0}$ (1.72-2.40)**	**< .001*****
Perceived usefulness of the app	-	-	**$137.32_{0,0}$ (49.45-381.34)**	**< .001*****
Alcohol consumption	-	-	-	-
Perceived lack of time	-	-	**$0.20_{0,0}$ (0.14-0.29)**	**< .001*****
P5				
Motivation to reduce alcohol	-	-	-	-
Perceived usefulness of the app	**$1.42_{2,2}$ (1.15-1.75)**	**< .01****	**$1.93_{0,0}$ (1.06-1.82)**	**.02***
Alcohol consumption	-	-	-	-
Perceived lack of time	$1.08_{2,2}$ (0.81-1.43)	.60	-	-
P6				
Reminder	**$3.88_{2,0}$ (1.37-11.03)**	**.01***	-	-
Motivation to reduce alcohol	$1.07_{2,0}$ (0.93-1.21)	.35	**$3.45_{0,0}$ (1.34-8.83)**	**.01***
Perceived usefulness of the app	$1.12_{2,0}$ (0.94-1.34)	.21	-	-
Alcohol consumption	$0.92_{2,0}$ (0.83-1.02)	.13	-	-
Perceived lack of time	**$0.77_{2,0}$ (0.61-0.97)**	**.03***	$1.24_{0,0}$ (0.71-2.17)	.45

Note. All models were adjusted for the day of the week using a cyclic cubic smoothing term. Numbers in subscript indicate the lags of autoregressive (AR) and moving average (MA) terms, respectively. A lag value of 0 indicates that an AR or MA term was not included; * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 7.5. *Continued.*

	Frequency of engagement		Amount of engagement	
	IRR (95% CI)	***p*-value**	**IRR (95% CI)**	***p*-value**
P7				
Reminder	**$3.26_{1,0}$ (2.15-4.96)**	**< .001*****	-	-
Motivation to reduce alcohol	-	-	**$1.67_{0,0}$ (1.16-2.40)**	**< .01****
Perceived usefulness of the app	-	-	**$0.52_{0,0}$ (0.33-0.80)**	**< .01****
Alcohol consumption	-	-	-	-
Perceived lack of time	-	-	-	-
P8				
Motivation to reduce alcohol	-	-	-	-
Perceived usefulness of the app	-	-	-	-
Alcohol consumption	$0.85_{1,0}$ (0.67-1.09)	.20	$0.82_{0,0}$ (0.47-1.43)	.50
Perceived lack of time	-	-	$1.33_{0,0}$ (0.97-1.82)	.08
P9				
Reminder	-	-	-	-
Motivation to reduce alcohol	-	-	$1.20_{1,1}$ (0.92-1.58)	0.18
Perceived usefulness of the app	**$1.38_{1,0}$ (1.24-1.53)**	**< .001*****	**$1.67_{1,1}$ (1.22-2.29)**	**< .01****
Alcohol consumption	-	-	-	-
Perceived lack of time	-	-	**$4.77_{1,1}$ (1.09-20.79)**	**.04***

Note. All models were adjusted for the day of the week using a cyclic cubic smoothing term. Numbers in subscript indicate the lags of autoregressive (AR) and moving average (MA) terms, respectively. A lag value of 0 indicates that an AR or MA term was not included; * $p < .05$; ** $p < .01$; *** $p < .001$.

7.5 Discussion

7.5.1 Summary of key findings

The current series of *N*-of-1 studies found that app-related and psychological variables identified as important for engagement with apps for alcohol reduction at the between-subjects level also fluctuate over time within individuals. The utility of these variables in predicting two distinct facets of behavioural engagement (i.e. the frequency and amount of engagement) with an alcohol reduction app, *Drink Less,* differed within and between individuals. This suggests that different strategies to promote engagement may be required for different individuals, and that such strategies may have differential effects on the various facets of engagement.

In line with findings from between-subjects studies [349], the receipt of a proactive reminder was significantly associated with the frequency of engagement for a few participants. However, this was not the case for all participants who had opted to have the reminder switched on during the study, which suggests that some participants may be more responsive to prompts than others. For some participants, the daily reminder was received in the middle of a measurement period. As all predictor variables were entered into the same model, it was not possible to assess whether the receipt of a reminder had a causal influence on subsequent engagement for these participants. Hence, for some participants (e.g. P3), the significant association between the receipt of a reminder and subsequent engagement may not indicate cause and effect.

In contrast to between-subjects studies [310], for a few participants, motivation to reduce alcohol was significantly associated with the amount, but not necessarily the frequency, of engagement with the *Drink Less* app. This may be interpreted to suggest that for some participants, being more highly motivated to reduce alcohol makes one more willing to spend time (and effort) on the app, providing that one has decided to open the app in the first place.

Previous between-subjects studies have identified a negative relationship of baseline alcohol consumption with the frequency of engagement, such that the higher the alcohol consumption, the less frequent the engagement. In the present study, none of the participants was found to engage with the app at a lower rate during time periods that followed sessions of heavier alcohol consumption. Instead, alcohol consumption was positively related to the frequency and amount of engagement for some participants. It is plausible that the directionality of the relationship between engagement and the target behaviour may vary across individuals: while some app users may be more prone to engage when they are 'doing well' (i.e. having abstained from or consumed less-than-typical amounts of alcohol), the reverse relationship may hold true for other users.

The variable 'perceived lack of time' has typically been explored qualitatively in interviews with participants who have dropped out of RCTs of DBCIs [164]. For some participants in the present study, this variable was significantly associated with fluctuations in the frequency and amount of engagement. However, the direction of the relationships varied across participants, with some participants displaying lower rates of the frequency or amount of engagement after having indicated that they had a lot of time available for the app. It is possible that more frequent EMAs would help detect a potential non-null relationship between 'perceived lack of time' and engagement for some participants: as only two measures per day were taken, participants may have rated themselves as having a lot of time for the app at the time of the measurement, but this might have changed a few hours later, thus interfering with their app use.

In line with findings from between-subjects studies [53,350], the variable 'perceived usefulness of the app' was found to be one of the most consistent predictors of both the frequency and amount of engagement with the *Drink Less* app. However, the direction of the associations differed across participants. Although 'perceived usefulness of the app' tended to be positively associated with the frequency and amount of engagement, the reverse held true for some participants. Again, this might be indicative of the need to capture this variable at a higher resolution (i.e. more frequent EMAs). Alternatively, this variable may have been subject to social desirability (e.g. participants not wanting to disappoint the researcher), or the question used to assess this variable might have been misinterpreted by some participants. This highlights the need for rigorous piloting of study measures to ensure that participants' interpretation is in line with the expected interpretation.

The finding that none of the assessed predictor variables was significantly associated with the frequency and/or amount of engagement for some participants begs the question as to what was driving engagement for these participants. One plausible explanation is that these participants established a habit or routine to engage with the app (e.g. every time they turned off their TV in the evening, they checked up on the *Drink Less* app). This could be tested in future research to assess whether some participants display more temporal regularity in their frequency of engagement than others. If this were indeed the case, other users could potentially be encouraged to establish routines to promote engagement with the app [351].

7.5.2 Strengths

To the author's knowledge, this was the first study to examine within-person predictors of the frequency and amount of engagement with an alcohol reduction app. The predictors tested in this study were selected based on prior evidence from between-person studies (reviewed in Chapter 2) and in-depth qualitative studies with potential users of alcohol reduction apps (reported in Chapters 3 and 6). Adherence to the twice-daily EMAs was high (0-16% missing data), and the automatic recording of the outcome variables in real-time ensured that participant burden and missing outcome data were minimised. This study provided initial evidence that it is acceptable to participants to gather data in this manner.

7.5.3 Limitations

Despite being conceptualised as a series of observational *N*-of-1 designs, participants engaged with an active intervention in addition to the study materials, which consisted of behaviour change techniques known to alter cognitions and behaviour (e.g. prompts, self-monitoring). It is therefore possible that both predictor and outcome variables were subject to non-random fluctuations, which were caused by participants' engagement with the intervention itself, or with the study materials. However, as engagement with DBCIs cannot be studied in isolation, without asking participants to engage with a given intervention, it was not possible to overcome this particular limitation.

The study sample was almost exclusively female. As men tend to exhibit more alcohol-related problems than women [329,330], it is unclear whether the same patterns of results would be observed in a more balanced or male-dominated sample. None of the participants dropped out of the study, thus indicating that they must all have been highly motivated to take part in the research. It is therefore possible that different patterns of results may be obtained in samples of less committed participants.

In order to keep participant burden to a minimum, the experiential facets of engagement during each login session were not assessed. This study was therefore unable to highlight potentially interesting relationships between the predictor variables and experiential engagement. Future research should test the feasibility of employing both time- and event-prompted EMAs, meaning that users are sent a prompt to respond to a few questions about their experiential engagement immediately after having engaged with the app. This would need to be piloted carefully to ensure feasibility given the additional burden on participants.

7.5.4 Avenues for future research

Qualitative pre- and post-study interviews were conducted with participants in the present study. Data analysis is planned to investigate explanations for the differences identified between participants, and the unexpected direction of the association between ‘perceived usefulness of the app’ and behavioural engagement observed for some participants.

7.5.5 Conclusions

This series of *N*-of-1 designs found that predictors of the frequency and amount of engagement with an alcohol reduction app differ across individuals. The most consistent predictor of both frequency and amount of engagement was perceived usefulness of the app.

7.5.6 Next steps

This was the final empirical study of the thesis. Chapter 8 brings together key findings from the systematic review and the five empirical studies conducted as part of this thesis, and highlights implications for research, policy and practice, in addition to avenues for future research.

8 CHAPTER 8 – General discussion

This thesis used an interdisciplinary approach, drawing on theoretical frameworks and methods from the behavioural science and HCI literatures, to study the problem of engagement with DBCIs, focusing on apps for smoking cessation and alcohol reduction. Chapters 2-7 of this thesis reported results from one systematic review and five empirical studies that used a range of qualitative and quantitative methods to address the following research objectives:

1. To gain a better understanding of how to conceptualise engagement with DBCIs
2. To gain a better understanding of how to measure engagement with DBCIs
3. To identify factors that promote or detract from engagement with DBCIs in general, and with smoking cessation and alcohol reduction apps in particular

In this final chapter, the key findings obtained in relation to the research objectives are first discussed. The following sections provide a reflection on general strengths and limitations of the research process, as more detailed issues pertaining to each study are covered at the end of each chapter. The final sections of this chapter focus on the implications for research, policy and practice, and unanswered questions and suggestions for future research arising from this thesis.

8.1 Summary and interpretation of key findings

8.1.1 Objective 1: To gain a better understanding of how to conceptualise engagement with DBCIs

Some form of engagement is necessary for DBCIs to be effective. Different conceptualisations of engagement have emerged both within and across scientific disciplines. Through i) integrating existing definitions of engagement from the behavioural science and HCI literatures identified in a systematic review (Chapter 2), ii) asking potential DBCI users how they understand the term 'engagement' (Chapter 4), and iii) the use of logical reasoning methods to identify necessary and sufficient conditions for someone to be engaged with a DBCI (Chapter 4), engagement with DBCIs was conceptualised here as a state that occurs during the momentary interaction with a DBCI. It was proposed that the state of engagement necessarily involves two behavioural (i.e. the amount and depth of DBCI use) and three experiential facets (i.e. attention, interest and enjoyment). This argument was partially supported by findings from two empirical studies (Chapters 4 and 5).

The practical utility of the proposed conceptualisation of engagement was examined by first constructing a self-report measure that assessed the five dimensions of engagement, subsequently examining how far these dimensions were related to one another and to key outcome variables in two empirical studies (Chapters 4 and 5). The results showed that engagement can be usefully defined both as a behaviour and as a subjective experience, and that this conceptualisation of engagement can be teased apart from other psychological states, such as motivation to change the target behaviour.

However, the hypothesis that engagement is underpinned by five (i.e. attention, interest, enjoyment, amount of use, depth of use) factors was not supported. Rather, evidence indicates that engagement is underpinned by two distinct factors, labelled ‘Experiential Engagement’ and ‘Behavioural Engagement’, respectively.

Theorists have argued that engagement with DBCIs includes cognitive dimensions not suggested in this thesis (e.g. the ability to comprehend the intervention materials and retain key information) [220], or that it does not include any experiential or cognitive dimensions beyond the behavioural dimensions (i.e. technology usage) [352]. As it is impossible to *objectively* determine the theoretical foundation of psychological constructs [353], the lack of consensus about the definition of engagement is to be expected. Even without this consensus, empirical tests of how key variables relate to one another, both initially and over a period of time, are critically important in the process of developing an operational definition of engagement that is useful for researchers, practitioners and policy-makers. This thesis found that the addition of the experiential indicators of engagement to a model including only the behavioural indicators led to an improvement in model fit when predicting subsequent behavioural engagement (reported in Chapters 4 and 5). This suggests that the definition of engagement proposed in this thesis has added predictive power compared with definitions focusing solely on technology usage.

8.1.2 Objective 2: To gain a better understanding of how to measure engagement with DBCIs

Existing ways of measuring engagement with DBCIs include objectively recorded usage data from apps and websites, self-report questionnaires, qualitative methods, observational methods, sensor data from wearables and psychophysical measures of attention and arousal (reviewed in Chapter 2). As it was considered useful to be able to compare engagement levels across DBCIs in a standardised way, a self-report scale was developed. The 'DBCI Engagement Scale' was constructed with input from potential users and evaluated in two different populations of users who were willing to download and explore the *Drink Less* app (reported in Chapters 4 and 5). The overall measure was found to be predictive of subsequent behavioural engagement in the second, but not the first, evaluation study. Criterion and divergent validity were not established, and the scale's internal consistency reliability was questionable. Although it must be concluded that the 'DBCI Engagement Scale' requires further refinement before it is ready for routine use, the scale has demonstrated potential as an instrument that could be useful for researchers, policy-makers and practitioners.

8.1.3 To identify factors that promote or detract from engagement with DBCIs in general, and with smoking cessation and alcohol reduction apps in particular

A conceptual framework of factors that promote or detract from engagement with DBCIs was developed in Chapter 2. Specific influences on engagement with apps for smoking cessation and alcohol reduction were explored in

Chapters 3, 6 and 7 through the use of think aloud and interview techniques, focus group and survey methodology and a series of *N*-of-1 designs. The last two empirical studies focused exclusively on apps for alcohol reduction.

The factors judged to be most important for the uptake of apps for smoking cessation and alcohol reduction were the immediate look and feel of the app (which included perceived ease of use and appealing aesthetics), ‘social proof’ and titles that appear realistic (Chapter 3). Some factors were initially mentioned as important for engagement (i.e. credibility, accuracy, familiarity), but were identified as being more important for the initial uptake of apps (as opposed to subsequent engagement) in a later study using a different methodology (reported in Chapter 6).

The design features expected to be most important for engagement by potential users were those that enhance their motivation to change the target behaviour, foster their beliefs about the perceived usefulness and relevance of the app, and spark their interest (Chapters 2, 3 and 6). The relative importance of these factors in predicting the frequency and amount of engagement with the *Drink Less* app was found to differ both within and between individuals (Chapter 7). However, the most consistent psychological predictor of the frequency and amount of engagement was perceived usefulness of the app. Specific design recommendations based on these findings are outlined in the section labelled ‘Implications for research, policy and practice’.

The finding that perceived usefulness of the app is a key predictor of engagement with apps for smoking cessation and alcohol reduction lends partial support to the Technology Acceptance Model (TAM) [52]. The TAM is centred on two constructs – perceived ease of use and perceived usefulness – which are expected to jointly contribute to intentions to use technology. However, as intentions do not always translate into action (known in the behaviour change literature as the 'intention-behaviour gap') [54], the ability of TAM to predict actual DBCI engagement was unclear prior to this thesis. In particular, the findings from the series of *N*-of-1 designs (Chapter 7) suggest that perceived usefulness is not only predictive of intentions to engage, but also of behavioural engagement. This finding, coupled with the observation that participants in Study 5 expected that design features that spark their interest are important for engagement, can also be interpreted within the context of Ryan and Deci's Self-Determination Theory (SDT). The SDT distinguishes between 'intrinsic motivation' (i.e. the performance of an activity for no apparent reason other than it being perceived as enjoyable or interesting in itself) and 'extrinsic motivation' (i.e. the performance of an activity because it is perceived to be instrumental in achieving some other valued outcome, distinct from the activity itself) [55]. The observation that users are driven to engage with apps when they perceive these to be useful for achieving a particular goal (e.g. reducing their drinking) suggests that extrinsic motivation plays an important role in the promotion of engagement with DBCIs. Given that DBCIs are designed to achieve outcomes that are distinct from the activity of engaging with the DBCI itself, this is not a surprising finding. However, the finding that potential users expected that they would be more prone to engage with a DBCI if it sparked their interest suggests that intrinsic motivation to engage with the DBCI itself also plays a key role in the promotion of engagement with DBCIs.

The results from this thesis can also be linked to the User Experience (UX) perspective, which is concerned with the ability of interactive products to satisfy users' need for autonomy (defined as the feeling of being in control of one's actions), stimulation (defined as the feeling of pleasure and interest), meaning, or relatedness to other people [56,57]. Many participants in the studies reported in Chapters 3 and 6 mentioned that they did not want to share progress or discuss behaviour change strategies with other app users. Coupled with the finding that perceived usefulness and perceived personal relevance were identified as key predictors of engagement, this may be interpreted to suggest that DBCI users not only have a need for relatedness to other people (which may or may not be mediated by the technology), but also a need for 'relatedness to the technology itself', which could perhaps be defined as 'the feeling that the app speaks directly to the user'.

The findings from this thesis also lend empirical support to the Behaviour Change Intervention Ontology [9]. Results from the studies reported in Chapters 2, 3, 6 and 7 indicate that the DBCI itself (i.e. content and delivery), the context of use (i.e. the setting in which the DBCI is used and characteristics of the population using it) and the target behaviour (e.g. alcohol consumption) do indeed influence engagement with DBCIs. The finding that not only baseline, but also daily levels of motivation to change and perceived usefulness of the app when interacting with a DBCI are significantly associated with engagement, may be interpreted to suggest that 'mechanisms of action' of the DBCI (i.e. psychological processes that change due to interactions with the intervention) also influence engagement.

8.2 Strengths

This thesis has several strengths, particularly from a methodological viewpoint: a range of qualitative and quantitative methods (i.e. mixed-methods) were used to address the same research questions. This helped overcome well-known limitations associated with each method, as data sources were triangulated. The use of novel research methods (e.g. the new ranking task paradigm developed as part of Chapter 6, the series of *N*-of-1 designs in Chapter 7) also facilitated this. Moreover, the research conducted as part of this thesis was interdisciplinary in scope, drawing on theoretical frameworks and methods from the behavioural science and HCI literatures. Although interdisciplinary research comes with its particular challenges [354], the interdisciplinary approach constitutes a key strength of this thesis, as a broader range of factors were considered, and a wider range of research methods were employed, than if the research had been confined to a single scientific discipline.

8.3 Limitations

This thesis had a broad scope: rather than focusing efforts either on the development of a self-report scale or on the identification of factors that promote engagement, both of these areas were covered. More work is required to refine and test the 'DBCI Engagement Scale' before it can be used in routine practice, and more experimental work is required to test the conceptual framework of factors that influence engagement with DBCIs.

Due to the nature of the methods used in this thesis, little knowledge was generated with regards to the specific characteristics of the setting of use that may influence engagement with apps for smoking cessation and alcohol reduction. Moreover, this thesis focused only on one mode of delivery (i.e. smartphone apps) and two different behaviours, with more emphasis on alcohol than smoking towards the end of the thesis. Although research methods similar to those used in this thesis could be employed to test whether the findings generalise to DBCIs for other behaviours (e.g. physical activity), different methods are required to identify what environmental factors may influence engagement with apps for smoking cessation and alcohol reduction. For example, the use of sensor data, such as the smartphone's global positioning system, could be used for this purpose.

8.4 Implications for research, policy and practice

8.4.1 Research

Due to the observed non-normal distributions of the scale items that jointly formed the 'DBCI Engagement Scale', a decision was made to use *z*-score normalisation. Consequently, total scores on the 'DBCI Engagement Scale' are only meaningful in relation to the average intensity of experiential and behavioural engagement that a particular DBCI generates. This may facilitate attempts to develop cut-offs for 'high' and 'low' engagers across DBCIs, irrespective of their specific parameters (e.g. the number and length of intervention components). For example, users with scores that fall within a particular range of standard deviations above or below the mean might usefully be classified as 'high' and 'low' engagers, respectively, and these patterns may replicate across DBCIs. This merits exploration by evaluating the 'DBCI Engagement Scale' across different kinds of DBCIs (e.g. apps for smoking cessation or physical activity).

Moreover, the finding that fluctuations in psychological and app-related variables are associated with variability in engagement suggests that different intervention strategies may be effective for different users, at different points in time. Hence, just-in-time adaptive interventions (JITAIs) [274] may be a promising intervention strategy to test in future research. The JITAI is a type of intervention that is specifically designed to address the dynamically changing needs of individuals. JITAIs rely on inputs from, for example, EMAs or sensor data collected via wearables to make decisions as to whether an intervention should be delivered at a particular moment in time or not and if so, what type of intervention to deliver. For example, a JITAI could be delivered when an individual's level of perceived usefulness of the app or motivation to reduce alcohol is below a given threshold for action. The utility of JITAIs for promoting engagement with apps for alcohol reduction should be explored in future research.

8.4.2 Policy

The results of this thesis have implications for digital health policy: a shared definition and measure of engagement can be used to help policy-makers and commissioners to set evaluation standards for health apps and other DBCIs. For example, the UK National Health Service's Apps Library (https://apps.beta.nhs.uk/), which endorses health apps that meet particular criteria, currently asks developers to provide evidence of effectiveness, usability, technical stability and data protection. Findings from this thesis suggest that the Apps Library should also require developers to report information about observed engagement levels in different subgroups of users, to ensure that products are of a certain standard before they are promoted more widely.

8.4.3 Practice

As smokers and drinkers tend to select apps at least partly based on their immediate look and feel, it is important for healthcare professionals to collaborate with interaction design experts to develop evidence-based smoking cessation and alcohol reduction apps that are on a par with other commercially available apps in terms of aesthetics and usability. As smokers and drinkers were found to rely on 'social proof' (i.e. other users' ratings and brand recognition) when selecting apps, this could be leveraged by researchers and practitioners by initiating collaborations with developers of popular apps or apps from well-known brands. For example, it might be more fruitful to modify the content of a well-established app with an existing client base rather than developing a novel smoking cessation or alcohol reduction app.

The finding that users may continue to engage with smoking cessation and alcohol reduction apps only if they are regularly provided with information or features that pique their interest suggests that this needs to be considered in the design process. The possibility of preventing disengagement by providing features that meet users' need for stimulation (e.g. novelty, narrative features, interactive features) should therefore be explored.

The finding that smokers and drinkers are more willing to engage with apps that provide options regarding quitting strategy poses a design challenge. As evidence suggests that some quitting strategies are more effective than others on average – for example, quitting smoking 'cold turkey' tends to be more effective than gradual reduction [355] – designers might benefit from using persuasive design elements (as suggested by the Persuasive Systems Design Model), such as providing tutorials and guidance, using tunnelling techniques (i.e. making users click through a pre-specified sequence of pages), or making use of normative influence, to attempt to modify users' beliefs and attitudes [61].

Findings from this thesis also suggest that the specifics of how to personalise content to support smokers' and drinkers' needs merit further investigation. A data-driven approach using machine-learning techniques might be helpful in advancing the knowledge on how to meaningfully tailor app content according to individual differences. For example, machine-learning techniques were recently applied to data from the 'E-COMPARED' project, an RCT comparing treatment as usual with blended therapy (i.e. internet-based therapy in combination with face-to-face support) in adults with major depressive disorder. Baseline data (e.g. depression and anxiety scores) were used to predict outcomes and treatment cost, which were subsequently used to derive individual treatment recommendations [356]. In a similar vein, decisions about the type of feedback to provide in apps for smoking cessation and alcohol reduction, or whether or not to offer features that link users with others on social media, could be made based on individual preferences at baseline, to foster a sense of perceived usefulness and personal relevance. It has been proposed that tailoring of content or features based on psychological constructs (e.g. the need for relatedness to other people or to the technology itself) is more effective than tailoring based on behaviour, which is in turn more effective than tailoring based on demographic characteristics [357]. Tailoring on users' underlying psychological needs, such as the need for relatedness or stimulation, thus constitutes an avenue for future research.

The finding that few smokers and drinkers wanted to use the apps in social settings should be considered in the design process. Smoking and drinking are perceived as more private than, for example, physical activity behaviours, perhaps due to social stigma [271,272]. It should therefore not be assumed that features typically included in apps targeting other types of behaviour (e.g. physical activity) can successfully be transferred to those targeting smoking and drinking. The hypothesis that smokers and drinkers might engage more with apps that suggest how to replace smoking and drinking with other activities, as opposed to those that provide in-the-moment support, could be tested in future research. See Table 8.1. for a summary of design recommendations.

Table 8.1. Summary of key design recommendations.

Category	*Design Recommendations*
How can the reach of evidence-based apps be improved?	Develop smoking cessation and alcohol reduction apps that are on a par with other commercially available apps in terms of aesthetics and usability, perhaps including collaboration with interaction design experts.
	Researchers and practitioners may consider initiating collaborations with developers of popular apps and/or apps from well-known brands to leverage their existing 'social proof'.
	Use simple and straightforward titles that include key words (e.g. "quit smoking" or "drink less alcohol").
How can engagement with apps for smoking cessation and alcohol reduction be improved?	Use persuasive design elements (e.g. guidance, tunnelling, personalisation, normative influence) to modify users' beliefs about how to quit smoking or reduce their drinking, or their beliefs about the app being personally relevant to them and their particular situation.
	Consider users' need for stimulation with a view to sparking users' interest during each DBCI interaction (e.g. novelty, interactivity, narrative features).
	Use machine-learning techniques to explore how to meaningfully tailor content according to individual differences.
	Consider the online and offline social preferences of the target population. For example, it might be more fruitful to focus on behaviour substitution or problem solving, as opposed to in-the-moment support, for smokers and drinkers.

8.5 Unanswered questions and avenues for future research

Engagement was conceptualised here as a state with experiential and behavioural facets which can vary in intensity. The claim that engagement is a state (rather than an enduring trait) was not tested empirically. State variables should vary as situational contingencies change (i.e. they should have low test-retest reliability) [290]. Future research should therefore test whether, when holding contextual factors constant (e.g. population and setting of use), the state of engagement varies across different DBCIs (e.g. two differently designed apps for alcohol reduction).

In line with theories of behaviour change [358], engagement with DBCIs may be more usefully conceptualised as a behaviour that is influenced by multiple, dynamically interacting intra- and extra- individual factors (e.g. psychological, social, environmental). It may therefore be more fruitful to consider how different configurations of intra- and extra-individual factors dynamically interact over short time-periods (e.g. hours, days) to influence behaviour (sometimes referred to as 'state-space representations' of when a particular intervention produces a given effect) [359]. For example, the likelihood that a user engages behaviourally with an alcohol reduction app (i.e. the frequency of engagement) may increase if, for example, (i) their belief that the app is useful to them is above a certain threshold for action, (ii) their daily level of motivation to reduce drinking is high, and (iii) they are not surrounded by others who drink. However, the likelihood that a user spends time on the app (conditional on them having decided to engage with the app in the first place) may increase if, for example, (i) they are experiencing enjoyment and/or interest whilst interacting with the app (also referred to as 'intrinsic motivation' [55]), (ii) they are not stressed, and (iii) they are not distracted by external stimuli. The inter-relationships between such variables should be tested using EMAs to gather temporally rich, contextualised data [226], which can be modelled using computational techniques from control systems engineering (e.g. dynamic systems modelling) [275].

Moreover, future attempts to validate state-based measures of engagement should carefully consider other indicators of predictive validity. For example, it is plausible that greater intensity of initial engagement predicts knowledge or skills at a future time point, as suggested by the Elaboration Likelihood Model of Persuasion (ELMP) [63]. The ELMP argues that deeper information processing (which tends to occur when one is paying attention to the content), which can be manipulated by persuasive design features, leads to better knowledge retention. This merits exploration in future research.

Findings from this thesis also beg the question as to what sources of motivation are most supportive of engagement. This should be investigated experimentally through A/B testing or factorial experiments. It may, for example, be hypothesised that features that support users' intrinsic motivation to engage with the app (e.g. novelty features, interactive features) will differentially impact on the frequency of engagement, as compared with features that support users' extrinsic motivation (e.g. perceived usefulness and relevance of the app).

8.6 Concluding remarks

Some form of engagement is necessary for DBCIs to be effective. This thesis aimed to gain a better understanding of how to define, measure and promote engagement with DBCIs in general, and with apps for smoking cessation and alcohol reduction in particular. This was achieved through the use of a range of qualitative and quantitative methods. Engagement was usefully defined both as a subjective experience and as a behaviour, and a self-report measure with promising psychometric properties was developed and evaluated. Results from one-to-one interviews, focus groups and a series of *N*-of-1 designs showed that design features that enhance users' motivation to change the target behaviour, that foster their beliefs about the perceived usefulness and relevance of the app, and that spark their interest are most important for promoting engagement with apps for smoking cessation and alcohol reduction. These findings can be used to inform the design of new, or modification of existing, apps for these behaviours.

References

1. Belloc NB, Breslow L. Relationship of Physical Health Status and Health Practices. Prev Med (Baltim) 1972;1:409–421.

2. Lancaster T, Stead LF. Individual behavioural counselling for smoking cessation. Cochrane Database Syst Rev 2005;(2):CD001292. PMID: 15846616

3. Kaner EFS, Dickinson HO, Beyer F, Pienaar E, Schlesinger C, Campbell F, et al. The effectiveness of brief alcohol interventions in primary care settings: A systematic review. Drug Alcohol Rev 2009;28(3):301–323. PMID: 19489992

4. Gordon L, Graves N, Hawkes A, Eakin E. A review of the cost-effectiveness of face-to-face behavioural interventions for smoking, physical activity, diet and alcohol. Chronic Illn [Internet] 2007;3(2):101–129. Available from: http://journals.sagepub.com/doi/10.1177/1742395307081732 PMID: 18083667

5. Action on Smoking and Health. Cutting Down: The Reality of Budget Cuts to Local Tobacco Control [Internet]. Cancer Res UK. 2016. Available from: http://www.cancerresearchuk.org/sites/default/files/local_authority_survey_2016_report_cruk_finalfinal.pdf

6. Brown J, West R, Angus C, Beard E, Brennan A, Drummond C, et al.

Comparison of brief interventions in primary care on smoking and excessive alcohol consumption: A population survey in England. Br J Gen Pract 2016;66(642):e1–e9.

7. Office for National Statistics. Internet access - households and individuals: 2016 [Internet]. 2016. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2015-08-06

8. Perrin A, Duggan M. Americans' Internet access: 2000-2015. Pew Res Cent. 2015.

9. West R, Michie S. A Guide to Development and Evaluation of Digital Interventions in Healthcare. London: Silverback Publishing; 2016.

10. Khadjesari Z, Stevenson F, Godfrey C, Murray E. Negotiating the "grey area between normal social drinking and being a smelly tramp": A qualitative study of people searching for help online to reduce their drinking. Heal Expect 2015;18:2011–2020. PMID: 25676536

11. Ritterband LM, Tate DF. The Science of Internet Interventions. Ann Behav Med 2009;38:1–3. PMID: 19816750

12. Schueller SM, Muñoz RF, Mohr DC. Realizing the Potential of Behavioral Intervention Technologies. Curr Dir Psychol Sci 2013;22(6):478–483.

13. Strecher V. Internet Methods for Delivering Behavioral and Health-

Related Interventions (eHealth). Annu Rev Clin Psychol [Internet] 2007;3:53–76. Available from: http://www.annualreviews.org/doi/10.1146/annurev.clinpsy.3.022806.091428 PMID: 17716048

14. Della Mea V. What is e-health (2): The death of telemedicine? J Med Internet Res 2001;3(2):6–7. PMID: 11720964

15. Taylor G, Dalili M, Semwal M, Civljak M, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2017;CD007078(9).

16. Kaner E, Beyer F, Garnett C, Crane D, Brown J, Muirhead C, et al. Personalised digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations. Cochrane Database Syst Rev [Internet] 2017;CD011479(9). Available from: http://doi.wiley.com/10.1002/14651858.CD011479

17. Whittaker R, Borland R, Bullen C, Rb L, Mcrobbie H, Rodgers A. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2009;(4).

18. Davies CA, Spence JC, Vandelanotte C, Caperchione C, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act [Internet] 2012 [cited 2015 Nov 6];9(52):1–13. Available from: http://www.biomedcentral.com/content/pdf/1479-5868-9-52.pdf

19. Liu F, Kong X, Cao J, Chen S, Li C, Huang J, et al. Mobile phone intervention and weight loss among overweight and obese adults: a meta-analysis of randomized controlled trials. Am J Epidemiol [Internet] 2015 Mar 1 [cited 2015 Dec 2];181(5):337–48. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25673817 PMID: 25673817

20. Muntaner A, Vidal-Conti J, Palou P. Increasing physical activity through mobile device interventions: A systematic review. Health Informatics J [Internet] 2015 Feb 3 [cited 2016 Feb 8];1–19. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25649783 PMID: 25649783

21. Jones KR, Lekhak N, Kaewluang N. Using mobile phones and short message service to deliver self-management interventions for chronic conditions: a meta-review. Worldviews Evidence-Based Nurs [Internet] 2014 Apr [cited 2016 Feb 8];11(2):81–88. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24597522 PMID: 24597522

22. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, et al. Effectiveness of E-Self-help Interventions for Curbing Adult Problem Drinking: A Meta-analysis. J Med Internet Res 2011;13(2):e42. PMID: 21719411

23. Riper H, Blankers M, Hadiwijaya H, Cunningham J, Clarke S. Effectiveness of Guided and Unguided Low-Intensity Internet Interventions for Adult Alcohol Misuse: A Meta-Analysis. PLoS One 2014;9(6):e99912.

24. Shahab L, McEwen A. Online support for smoking cessation: A

systematic review of the literature. Addiction 2009;104(11):1792–1804. PMID: 19832783

25. Webb TL, Joseph J, Yardley L, Michie S. Using the Internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4. PMID: 20164043

26. van Genugten L, Dusseldorp E, Webb TL, Van Empelen P. Which Combinations of Techniques and Modes of Delivery in Internet-Based Interventions Effectively Change Health Behavior? A Meta-Analysis. J Med Internet Res 2016;18(6):e155. PMID: 27268104

27. Eysenbach G. The law of attrition. J Med Internet Res [Internet] 2005 Jan [cited 2015 Feb 10];7(1):e11. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1550631&tool=pmcentrez&rendertype=abstract PMID: 15829473

28. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res [Internet] Journal of Medical Internet Research; 2012 Jan 14 [cited 2015 Oct 28];14(6):e152. Available from: http://www.jmir.org/2012/6/e152/ PMID: 23151820

29. Appboy. Spring 2016 Mobile Customer Retention Report: An Analysis of Retention by Day. 2016.

30. Consumer Health Information Corporation. Motivating Patients to Use Smartphone Health Apps [Internet]. 2015 [cited 2015 Aug 10]. Available from: http://www.prweb.com/releases/2011/04/prweb5268884.htm

31. Cobb NK, Graham AL, Bock BC, Papandonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. Nicotine Tob Res 2005;7(2):207–216. PMID: 16036277

32. Richardson A, Graham AL, Cobb N, Xiao H, Mushro A, Abrams D, et al. Engagement promotes abstinence in a web-based cessation intervention: cohort study. J Med Internet Res [Internet] Journal of Medical Internet Research; 2013 Jan 28 [cited 2016 Jan 19];15(1):e14. Available from: http://www.jmir.org/2013/1/e14/ PMID: 23353649

33. Alexander GL, McClure JB, Calvi JH, Divine GW, Stopponi MA, Rolnick SJ, et al. A randomized clinical trial evaluating online interventions to improve fruit and vegetable consumption. Am J Public Health 2010;100(2):319–326. PMID: 20019315

34. Tate DF, Wing RR, Winett R a. Using Internet technology to deliver a behavioral weight loss program. J Am Med Assoc 2001;285(9):1172–1177. PMID: 11231746

35. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. Am J Prev Med [Internet] Elsevier; 2016;51(5):833–842. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0749379716302434

36. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res [Internet] 2011 Jan [cited 2015 Oct 5];13(3):e52. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3222162&tool=pmcentrez&rendertype=abstract PMID: 21821503

37. Graham AL, Stanton CA, Papandonatos GD, Erar B. Use of an online smoking cessation community promotes abstinence: Results of propensity score weighting. Heal Psychol 2015;34:1286–1295. PMID: 26651470

38. Pagliari C. Design and evaluation in ehealth: Challenges and implications for an interdisciplinary field. J Med Internet Res 2007;9(2):e15. PMID: 17537718

39. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2008 Jan 4 [cited 2016 Feb 8];10(5):e36. Available from: http://www.jmir.org/2008/5/e36/ PMID: 18984557

40. Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in Web-based health behavior change programs. J Med Internet Res [Internet] 2006 Jan 30 [cited 2015 Mar 29];8(3):e15. Available from: http://www.jmir.org/2006/3/e15/ PMID: 16954125

41. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: A model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011;13(1):e30. PMID: 21393123

42. Donkin L, Hickie IB, Christensen H, Naismith SL, Neal B, Cockayne NL, et al. Rethinking the dose-response relationship between usage and outcome in an online intervention for depression: Randomized controlled trial. J Med Internet Res 2013;15(10):e231. PMID: 24135213

43. Csikszentmihalyi M. Flow: The psychology of optimal performance. New York: Cambridge University Press; 1990.

44. Jennett C, Cox AL, Cairns P, Dhoparee S, Epps A, Tijs T, et al. Measuring and Defining the Experience of Immersion in Games. Int J Hum Comput Stud 2008;66(9):641–661.

45. Brown E, Cairns P. A grounded investigation of game immersion. CHI ’04 Ext Abstr Hum Factors Comput Syst [Internet] ACM; 2004. p. 1297–1300. Available from: http://doi.acm.org/10.1145/985921.986048

46. O’Brien HL, Toms EG. The Development and Evaluation of a Survey to Measure User Engagement. J Am Soc Inf Sci Technol 2010;61(1):50–69. PMID: 502955140

47. Jackson SA, Marsh HW. Development and validation of a scale to measure optimal experience: The Flow State Scale. J Sport Exerc Psychol [Internet] 1996;18:17–35. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=s3h&AN=960409

1722&site=ehost-live PMID: 9604091722

48. Michie S, West R, Campbell R, Brown J, Gainforth H. ABC of behaviour change theories. Silverback Publishing; 2014.

49. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implement Sci [Internet] BioMed Central Ltd; 2011;6(1):42. Available from: http://www.implementationscience.com/content/6/1/42 PMID: 21513547

50. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013;46(1):81–95. PMID: 23512568

51. Michie S, Thomas J, Johnston M, Aonghusa P Mac, Shawe-Taylor J, Kelly MP, et al. The Human Behaviour-Change Project: Harnessing the power of artificial intelligence and machine learning for evidence synthesis and interpretation. Implement Sci Implementation Science; 2017;12(1):1–12. PMID: 29047393

52. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Q 1989;13(3):319–340.

53. King WR, He J. A meta-analysis of the technology acceptance model. Inf Manag 2006;43(6):740–755. PMID: 25246403

54. Armitage CJ, Conner M. Efficacy of the Theory of Planned Behaviour: A meta-analytic review. Br J Soc Psychol [Internet] 2001 Dec;40:471–99. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11795063 PMID: 11795063

55. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55(1):68–78. PMID: 11392867

56. Hassenzahl M, Tractinsky N. User experience - a research agenda. Behav Inf Technol 2006;25(2):91–97.

57. Hassenzahl M, Diefenbach S, Göritz A. Needs, affect, and interactive products - Facets of user experience. Interact Comput [Internet] Elsevier B.V.; 2010;22(5):353–362. Available from: http://dx.doi.org/10.1016/j.intcom.2010.04.002

58. Law EL-C, Roto V, Hassenzahl M, Vermeeren APOS, Kort J. Understanding, scoping and defining user experience. Proc 27th Int Conf Hum Factors Comput Syst - CHI ’09 [Internet] 2009. p. 719. Available from: http://dl.acm.org/citation.cfm?doid=1518701.1518813 PMID: 2175

59. Hassenzahl M. The Interplay of Beauty, Goodness, and Usability in Interactive Products. Human-Computer Interact 2004;19(4):319–349.

60. Hassenzahl M, Monk A. The Inference of Perceived Usability From Beauty. Human-Computer Interact [Internet] 2010;25(3):37–41. Available from: http://dx.doi.org/10.1080/073700242010500139

61. Oinas-Kukkonen H, Harjumaa M. Persuasive Systems Design: Key Issues, Process Model, and System Features. Commun Assoc Inf Syst 2009;24(28):486–501.

62. Ludden GD, van Rompay TJ, Kelders SM, van Gemert-Pijnen JE. How to Increase Reach and Adherence of Web-Based Interventions: A Design Research Viewpoint. J Med Internet Res [Internet] 2015 Jul 10 [cited 2015 Jul 14];17(7):e172. Available from: http://www.jmir.org/2015/7/e172/

63. Petty RE, Cacioppo J. The Elaboration Likelihood Model of Persuasion. Adv Exp Soc Psychol 1986;19:123–205.

64. Kelders SM, Pots WTM, Oskam MJ, Bohlmeijer ET, Gemert-Pijnen JEWC Van. Development of a web-based intervention for the indicated prevention of depression. BMC Med Informatics Decis 2013;13(26):1–11.

65. Vines J, Clarke R, Wright P, McCarthy J, Olivier P. Configuring Participation: On How We Involve People In Design. CHI '13 Chang Perspect 2013. p. 429–438.

66. International Standardisation Organisation. ISO 9241-210. Ergonomics of Human-Computer Interaction - Part 2010: Human-Centred Design for Interactive Systems. 2010.

67. DeSmet A, Thompson D, Baranowski T, Palmeira A, Verloigne M, De Bourdeaudhuij I. Is Participatory Design Associated with the Effectiveness of Serious Digital Games for Healthy Lifestyle Promotion? A Meta-Analysis. J Med Internet Res 2016;18(4):e94.

68. Band R, Bradbury K, Morton K, May C, Michie S, Mair FS, et al. Intervention planning for a digital intervention for self-management of hypertension: a theory- , evidence- and person-based approach. Implement Sci Implementation Science; 2017;12(25):1–13.

69. Yardley L, Ainsworth B, Arden-Close E, Muller I. The person-based approach to enhancing the acceptability and feasibility of interventions. Pilot Feasibility Stud [Internet] Pilot and Feasibility Studies; 2015;37(1):1–7. Available from: http://dx.doi.org/10.1186/s40814-015-0033-z

70. DeSmet A, Palmeira A, Beltran A, Brand L, Fernandes Davies V, Thompson D. The Yin and Yang of Formative Research in Designing Serious (Exer-)games. Games Health J 2015;4(1):1–4.

71. Blandford A, Furniss D, Makri S. Qualitative HCI Research: Going Behind the Scenes. Synth Lect Human-Centered Informatics Morgan Claypool; 2016;9(1):1–115.

72. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012;380(9859):2224–2260. PMID: 23245609

73. Fidler JA, Shahab L, West O, Jarvis MJ, McEwen A, Stapleton JA, et al. “The smoking toolkit study”: a national study of smoking and smoking cessation in England. BMC Public Health [Internet] BioMed Central Ltd; 2011;11(1):479. Available from: http://www.biomedcentral.com/1471-

2458/11/479 PMID: 21682915

74. Gowing LR, Ali RL, Allsop S, Marsden J, Turf EE, West R, et al. Global statistics on addictive behaviours: 2014 status report. Addiction 2015;110:904–919.

75. Saunders JB, Aasland OG, Babor TF, Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption - II. Addiction 1993;88:791–804.

76. Department of Health. UK Chief Medical Officers' Alcohol Guidelines Review: Summary of the proposed new guidelines. 2016.

77. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care. 2nd ed. Geneva: World Health Organisation; 2001.

78. Beaglehole R, Bonita R. Alcohol: a global health priority. Lancet [Internet] Elsevier Ltd; 2009;373(9682):2173–2174. Available from: http://dx.doi.org/10.1016/S0140-6736(09)61168-5 PMID: 19560583

79. Myung S-K, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and Computer-Based Smoking Cessation Programs. Arch Intern Med 2009;169(10):929–937. PMID: 41045170

80. West R. The multiple facets of cigarette addiction and what they mean for encouraging and helping smokers to stop. COPD 2009;6(4):277–283.

PMID: 19811387

81. West R, Brown J. Theory of Addiction. 2nd ed. Wiley-Blackwell; 2013.

82. Bricker JB, Mull KE, Kientz J a, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend [Internet] Elsevier Ireland Ltd; 2014;143:87–94. Available from: http://www.sciencedirect.com/science/article/pii/S0376871614009776%5Cnhttp://ac.els-cdn.com/S0376871614009776/1-s2.0-S0376871614009776-main.pdf?_tid=f0b50af8-2c03-11e4-a7dd-00000aab0f27&acdnat=1408935796_46dd404812cc13cda580e703c6084b67 PMID: 25085225

83. Spruijt-Metz D, Hekler E, Saranummi N, Intille S, Korhonen I, Nilsen W, et al. Building new computational models to support health behavior change and maintenance: new opportunities in behavioral research. Transl Behav Med 2015;5(3):335–346.

84. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med [Internet] Elsevier; 2013;45(6):732–736. Available from: http://dx.doi.org/10.1016/j.amepre.2013.07.008 PMID: 24237915

85. Weaver ER, Horyniak DR, Jenkinson R, Dietze P, Lim MSC. "Let's get Wasted!" and Other Apps: Characteristics, Acceptability, and Use of Alcohol-Related Smartphone Applications. J Med Internet Res 2013;15(6):1–13. PMID: 25100681

86. Crane D, Garnett C, Brown J, West R, Michie S. Behavior change techniques in popular alcohol reduction apps: content analysis. J Med Internet Res [Internet] 2015;17(5):e118. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4468601&tool=pmcentrez&rendertype=abstract PMID: 25977135

87. Hoeppner BB, Hoeppner SS, Seaboyer L, Schick MR, Wu GWY, Bergman BG, et al. How Smart are Smartphone Apps for Smoking Cessation? A Content Analysis. Nicotine Tob Res 2016;18(5):1025–1031. PMID: 26045249

88. Ubhi HK, Kotz D, Michie S, van Schayck OCP, Sheard D, Selladurai A, et al. Comparative analysis of smoking cessation smartphone applications available in 2012 versus 2014. Addict Behav [Internet] The Authors; 2016;58:175–181. Available from: http://dx.doi.org/10.1016/j.addbeh.2016.02.026 PMID: 26950256

89. Choi J, Noh GY, Park DJ. Smoking cessation apps for smartphones: Content analysis with the self-determination theory. J Med Internet Res 2014;16(2):1–19. PMID: 24521881

90. Cohn, A. M., Hunter-Reel, D., Hagman, B. T., & Mitchell J. Promoting Behavior Change from Alcohol Use through Mobile Technology: The Future of Ecological Momentary Assessment. Alcohol Clin Exp Res [Internet] 2011;35(12):2209–2215. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3221771/

91. Buller DB, Borland R, Bettinghaus EP, Shane JH, Zimmerman DE.

Randomized trial of a smartphone mobile application compared to text messaging to support smoking cessation. Telemed e-Health [Internet] 2014;20(3):206–214. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24350804 PMID: 24350804

92. Gustafson D, McTavish F, Chih M-Y, Atwood A, Johnson R, Boyle M. A smartphone application to support recovery from alcoholism: A randomized controlled trial. JAMA Psychiatry [Internet] 2014;71(5):566–572. Available from: http://archpsyc.jamanetwork.com/article.aspx?articleid=1847578

93. Dulin PL, Gonzalez VM, Campbell K. Results of a Pilot Test of a Self-Administered Smartphone-Based Treatment System for Alcohol Use Disorders: Usability and Early Outcomes. Subst Abus 2014;35(2):168–175.

94. Gajecki M, Berman A, Sinadinovic K, Rosendahl I, Andersson C. Mobile phone brief intervention applications for risky alcohol use among university students: a randomized controlled study. Addict Sci Clin Pract [Internet] 2014;9(11):1–12. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed12&NEWS=N&AN=24985342 PMID: 24985342

95. Crane D, Garnett C, Michie S, West R, Brown J. A smartphone app to reduce excessive alcohol consumption: Identifying the effectiveness of intervention components in a factorial randomised control trial. Sci Rep 2018;8(4384):1–11.

96. BinDhim NF, McGeechan K, Trevena L. Smartphone Smoking Cessation Application (SSC App) trial: a multicountry double-blind automated randomised controlled trial of a smoking cessation decision-aid “app”. BMJ Open 2018;8(1):e017105. PMID: 29358418

97. Garnett C, Crane D, West R, Brown J, Michie S. The development of Drink Less: an alcohol reduction smartphone app for excessive drinkers. Transl Behav Med 2018;1–12.

98. Garnett C, Crane D, West R, Brown J, Michie S. Identification of Behavior Change Techniques and Engagement Strategies to Design a Smartphone App to Reduce Alcohol Consumption Using a Formal Consensus Method. JMIR mHealth uHealth [Internet] 2015;3(2):e73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26123578 PMID: 26123578

99. Murray E, White IR, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: Adherence and retention in a web-based alcohol trial. J Med Internet Res 2013;15(8):1–11. PMID: 23996958

100. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. [Updated March 2011]. [Internet]. Higgins J, Green S, editors. 2011 [cited 2015 Nov 15]. Available from: www.cochrane-handbook.org

101. Krishnan A. What are Academic Disciplines? NCRM Work Pap Ser ESRC Natl Cent Res Methods. 2009.

102. Krishnan A. Five Strategies for Practising Interdisciplinarity [Internet]. NCRM Work Pap Ser ESRC Natl Cent Res Methods. 2009. Available from: http://eprints.ncrm.ac.uk/782/1/strategies_for_practising_interdisciplinarity.pdf

103. Couper MP, Alexander GL, Zhang N, Little RJA, Maddy N, Nowak MA, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. J Med Internet Res [Internet] Journal of Medical Internet Research; 2010 Jan 18 [cited 2015 Dec 14];12(4):e52. Available from: http://www.jmir.org/2010/4/e52/ PMID: 21087922

104. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health [Internet] 2009;30:273–92. Available from: http://www.annualreviews.org/doi/abs/10.1146/annurev.publhealth.031308.100235 PMID: 19296777

105. Brouwer W, Oenema A, Raat H, Crutzen R, De Nooijer J, De Vries NK, et al. Characteristics of visitors and revisitors to an Internet-delivered computer-tailored lifestyle intervention implemented for use by the general public. Health Educ Res 2010;25(4):585–595. PMID: 19897515

106. Schubart JR, Stuckey HL, Ganeshamoorthy A, Sciamanna CN. Chronic Health Conditions and Internet Behavioral Interventions: A Review of Factors to Enhance User Engagement. Comput Informatics, Nurs 2011;29(2):81–92.

107. Huberman MA, Miles MB. Data Management and Analysis Methods. Handb Qual Res Thousand Oaks: SAGE Publications; 1994. p. 428–443.

108. O'Brien HL, Toms EG. What is User Engagement? A Conceptual Framework for Defining User Engagement with Technology. J Am Soc Inf Sci Technol 2008;59(6):938–955.

109. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick L a. A behavior change model for internet interventions. Ann Behav Med 2009;38:18–27. PMID: 19802647

110. Short CE, Rebar AL, Plotnikoff RC, Vandelanotte C. Designing engaging online behaviour change interventions: A proposed model of user engagement. Eur Heal Psychol 2015;17(1):32–38.

111. Centre for Reviews and Dissemination U of Y. Systematic Reviews: CRD's guidance for undertaking reviews in healthcare [Internet]. Khan K, Riet G Ter, Glanville J, Sowden A, Kleijnen J, editors. 2008 [cited 2016 Feb 8]. Available from: http://eprints.whiterose.ac.uk/1139/

112. Dixon-Woods M, Cavers D, Agarwal S, Annandale E, Arthur A, Harvey J, et al. Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. BMC Med Res Methodol [Internet] 2006 Jan [cited 2015 Aug 12];6:35. Available from: http://www.biomedcentral.com/1471-2288/6/35 PMID: 16872487

113. Dixon-Woods M, Bonas S, Booth A. How can systematic reviews incorporate qualitative research? A critical perspective. Qual Res

[Internet] 2006;6(1):27–44. Available from: http://qrj.sagepub.com/content/6/1/27.short

114. Entwistle V, Firnigl D, Ryan M, Francis J, Kinghorn P. Which experiences of health care delivery matter to service users and why? A critical interpretive synthesis and conceptual map. J Health Serv Res Policy [Internet] 2012;17(2):70–78. Available from: http://hsr.sagepub.com/content/17/2/70.full PMID: 21967821

115. Kazimierczak KA, Skea ZC, Dixon-Woods M, Entwistle VA, Feldman-Stewart D, N'Dow JMO, et al. Provision of cancer information as a "support for navigating the knowledge landscape": Findings from a critical interpretive literature synthesis. Eur J Oncol Nurs [Internet] Elsevier Ltd; 2013;17(3):360–369. Available from: http://dx.doi.org/10.1016/j.ejon.2012.10.002 PMID: 23164925

116. Morrison L, Yardley L, Powell J, Michie S. What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. Telemed e-Health [Internet] 2012;18(2):137–44. Available from: http://online.liebertpub.com/doi/abs/10.1089/tmj.2011.0062 PMID: 22381060

117. Anderson PJ. Assessment and Development of Executive Function (EF) During Childhood. Child Neuropsychol 2002;8(2):71–82. PMID: 12638061

118. Thomson Reuters. EndNote X7. Philadelphia, USA; 2013.

119. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol 1993;46(5):423–429. PMID: 8501467

120. Dixon-Woods M, Sutton A, Shaw R, Miller T, Smith J, Young B, et al. Appraising qualitative research for inclusion in systematic reviews: a quantitative and qualitative comparison of three methods. J Health Serv Res Policy 2007;12(1):42–47.

121. Barbour RS. Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? Br Med J 2001;322:1115–1117. PMID: 11337448

122. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med [Internet] 2009 Jul 21 [cited 2014 Jul 9];6(7):e1000097. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2707599&tool=pmcentrez&rendertype=abstract PMID: 19621072

123. Bianchi-Berthouze N, Kim WW, Patel D. Does Body Movement Engage You More in Digital Game Play? And Why? Proc Int Conf Affect Comput Intell Interact 2007. p. 102–113.

124. Chou JC, Hung C, Hung Y. Design Factors of Mobile Games for Increasing Gamers' Flow Experiences. Proc 2014 IEEE ICMIT 2014. p. 137–139.

125. Sharek D, Wiebe E. Measuring Video Game Engagement Through the

Cognitive and Affective Dimensions. Simul Gaming 2014;45:569–592.

126. Zhou T. Understanding the effect of flow on user adoption of mobile games. Pers Ubiquitous Comput 2013;17:741–748.

127. Oh J, Sundar SS. How Does Interactivity Persuade? An Experimental Test of Interactivity on Cognitive Absorption, Elaboration, and Attitudes. J Commun 2015;65:213–236.

128. Bouvier P, Lavoue E, Sehaba K. Defining Engagement and Characterizing Engaged-Behaviors in Digital Gaming. Simul Gaming [Internet] 2014;45(4–5):491–507. Available from: http://sag.sagepub.com/cgi/doi/10.1177/1046878114553571

129. Schønau-Fog H, Bjørner T. “Sure, I Would Like to Continue”: A Method for Mapping the Experience of Engagement in Video Games. Bull Sci Technol Soc 2012;32(5):405–412.

130. McClure JB, Shortreed SM, Bogart A, Derry H, Riggs K, St John J, et al. The effect of program design on engagement with an internet-based smoking intervention: randomized factorial trial. J Med Internet Res [Internet] 2013 Jan [cited 2015 Mar 29];15(3):e69. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3636802&tool=pmcentrez&rendertype=abstract PMID: 23529377

131. Voils CI, King HA, Maciejewski ML, Allen KD, Jr WSY, Shaffer JA. Approaches for Informing Optimal Dose of Behavioral Interventions. Ann Behav Med 2014;48:392–401.

132. Wang J, Sereika SM, Chasens ER, Ewing LJ, Matthews JT, Burke LE. Effect of adherence to self-monitoring of diet and physical activity on weight loss in a technology-supported behavioral intervention. Patient Prefer Adherence 2012;6:221–226.

133. Calleja G. Digital Game Involvement. Games Cult 2007;2(3):236–260.

134. Lin JC-C. Online stickiness: its antecedents and effect on purchasing intention. Behav Inf Technol 2007;26(6):507–516.

135. Han JY, Kim J-H, Yoon HJ, Shim M, McTavish FM, Gustafson DH. Social and Psychological Determinants of Levels of Engagement with an Online Breast Cancer Support Group: Posters, Lurkers, and Non-Users. J Health Commun 2012;17(3):356–371.

136. Morrison L, Moss-Morris R, Michie S, Yardley L. Optimizing engagement with Internet-based health behaviour change interventions: Comparison of self-assessment with and without tailored feedback using a mixed methods approach. Br J Health Psychol 2014;19:839–855.

137. Burns CG, Fairclough SH. Use of auditory event-related potentials to measure immersion during a computer game. Int J Hum Comput Stud [Internet] Elsevier; 2015;73:107–114. Available from: http://dx.doi.org/10.1016/j.ijhcs.2014.09.002

138. Chiang Y-T, Lin SSJ, Cheng C-Y, Liu EZ-F. Exploring online game players' flow experiences and positive affect. Turkish Online J Educ Technol 2011;10(1):106–114.

139. Chung J, Gardner HJ. Temporal Presence Variation in Immersive Computer Games. Int J Hum Comput Interact 2012;28(8):511–529.

140. Fang X, Zhang J, Chan SS. Development of an Instrument for Studying Flow in Computer Game Play. Int J Hum Comput Interact 2013;29(7):456–470.

141. Harmat L, Manzano Ö De, Theorell T, Högman L, Fischer H, Ullén F. Physiological correlates of the flow experience during computer game playing. Int J Psychophysiol [Internet] Elsevier B.V.; 2015;97:1–7. Available from: http://dx.doi.org/10.1016/j.ijpsycho.2015.05.001

142. Hilvert-Bruce Z, Rossouw PJ, Wong N, Sunderland M, Andrews G. Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depressive disorders. Behav Res Ther [Internet] 2012 Aug [cited 2015 Nov 2];50(7–8):463–468. Available from: http://www.sciencedirect.com/science/article/pii/S0005796712000708

143. Lefebvre RC, Tada Y, Hilfiker SW, Baur C. The Assessment of User Engagement with eHealth Content: The eHealth Engagement Scale. J Comput Commun 2010;15:666–681.

144. Martey RM, Kenski K, Folkestad J, Feldman L, Gordis E, Shaw A, et al. Measuring Game Engagement: Multiple Methods and Construct Complexity. Simul Gaming 2014;45:528–547.

145. Bossen D, Buskermolen M, Veenhof C, de Bakker D, Dekker J.

Adherence to a web-based physical activity intervention for patients with knee and/or hip osteoarthritis: a mixed method study. J Med Internet Res [Internet] Journal of Medical Internet Research; 2013 Jan 16 [cited 2016 Feb 5];15(10):e223. Available from: http://www.jmir.org/2013/10/e223/ PMID: 24132044

146. Geraghty AWA, Torres LD, Leykin YAN, Mun RF. Understanding attrition from international internet health interventions: a step towards global eHealth. Health Promot Int 2012;28(3):442–452.

147. Cussler EC, Teixeira PJ, Going SB, Houtkooper LB, Metcalfe LL, Blew RM, et al. Maintenance of weight loss in overweight middle-aged women through the Internet. Obesity [Internet] 2008 May [cited 2015 Nov 3];16(5):1052–1060. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18309301 PMID: 18309301

148. Davies C, Corry K, Van Itallie A, Vandelanotte C, Caperchione C, Mummery WK. Prospective associations between intervention components and website engagement in a publicly available physical activity website: the case of 10,000 Steps Australia. J Med Internet Res [Internet] Journal of Medical Internet Research; 2012 Jan 11 [cited 2016 Feb 5];14(1):e4. Available from: http://www.jmir.org/2012/1/e4/ PMID: 22260810

149. Dennison L, Morrison L, Lloyd S, Phillips D, Stuart B, Williams S, et al. Does Brief Telephone Support Improve Engagement With a Web-Based Weight Management Intervention? Randomized Controlled Trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2014 Mar 28

[cited 2015 Nov 2];16(3):e95. Available from: http://www.jmir.org/2014/3/e95/

150. Glasgow RE, Christiansen SM, Kurz D, King DK, Woolley T, Faber AJ, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. J Med Internet Res [Internet] Journal of Medical Internet Research; 2011 Jan 25 [cited 2016 Jan 5];13(1):e9. Available from: http://www.jmir.org/2011/1/e9/ PMID: 21371992

151. Manwaring JL, Bryson SW, Goldschmidt AB, Winzelberg AJ, Luce KH, Wilfley DE, et al. Do Adherence Variables Predict Outcome in an Online Program for the Prevention of Eating Disorders? J Consult Clin Psychol 2008;76(2):341–346.

152. Morrison C, Doherty G. Analyzing Engagement in a Web-Based Intervention Platform Through Visualizing Log-Data. J Med Internet Res [Internet] Journal of Medical Internet Research; 2014 Nov 13 [cited 2016 Jan 9];16(11):e252. Available from: http://www.jmir.org/2014/11/e252/

153. Murray E, White IR, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention in a web-based alcohol trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2013 Jan 30 [cited 2016 Feb 8];15(8):e162. Available from: http://www.jmir.org/2013/8/e162/ PMID: 23996958

154. Poirier J, Cobb NK. Social influence as a driver of engagement in a web-based health intervention. J Med Internet Res [Internet] Journal of Medical Internet Research; 2012 Jan 22 [cited 2016 Feb 8];14(1):e36.

Available from: http://www.jmir.org/2012/1/e36/ PMID: 22356829

155. Arden-Close EJ, Smith E, Bradbury K, Morrison L, Dennison L, Michaelides D, et al. A Visualization Tool to Analyse Usage of Web-Based Interventions: The Example of Positive Online Weight Reduction (POWeR). J Med Internet Res [Internet] 2015 May 19 [cited 2015 Aug 10];2(1):e8. Available from: http://humanfactors.jmir.org/2015/1/e8/

156. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2013 Jan 15 [cited 2015 Feb 17];15(4):e32. Available from: http://www.jmir.org/2013/4/e32/ PMID: 23587561

157. Chen Z, Koh PW, Ritter PL, Lorig K, Bantum EOC, Saria S. Dissecting an Online Intervention for Cancer Survivors: Four Exploratory Analyses of Internet Engagement and Its Effects on Health Status and Health Behaviors. Heal Educ Behav 2015;42(1):32–45.

158. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res [Internet] Journal of Medical Internet Research; 2009 Jan 24 [cited 2016 Feb 3];11(2):e13. Available from: http://www.jmir.org/2009/2/e13/ PMID: 19403466

159. Crutzen R, Cyr D, de Vries NK. The role of user control in adherence to and knowledge gained from a website: randomized comparison between a tunneled version and a freedom-of-choice version. J Med Internet Res [Internet] Journal of Medical Internet Research; 2012 Jan 9 [cited 2016

Feb 5];14(2):e45. Available from: http://www.jmir.org/2012/2/e45/ PMID: 22532074

160. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res [Internet] Journal of Medical Internet Research; 2011 Jan 14 [cited 2016 Jan 9];13(1):e17. Available from: http://www.jmir.org/2011/1/e17/ PMID: 21320854

161. Henshaw H, McCormack A, Ferguson MA. Intrinsic and extrinsic motivation is associated with computer-based auditory training uptake, engagement, and adherence for people with hearing loss. Front Psychol 2015;6:1–13.

162. Hsu C-L, Lu H-P. Why do people play on-line games? An extended TAM with social influences and flow experience. Inf Manag 2004;41:853–868.

163. McCabe MP, Price E. Attrition From an Internet-Based Psychological Intervention for Erectile Dysfunction: Who is Likely to Drop Out? J Sex Marital Ther 2009;35(5):391–401.

164. Postel MG, de Haan HA, ter Huurne ED, van der Palen J, Becker ES, de Jong CAJ. Attrition in web-based treatment for problem drinkers. J Med Internet Res [Internet] Journal of Medical Internet Research; 2011 Jan 27 [cited 2016 Feb 8];13(4):e117. Available from: http://www.jmir.org/2011/4/e117/ PMID: 22201703

165. Johansson O, Michel T, Andersson G, Paxling B. Experiences of non-adherence to Internet-delivered cognitive behavior therapy: A qualitative study. Internet Interv [Internet] 2015 May [cited 2015 Mar 23];2:137–142. Available from: http://www.sciencedirect.com/science/article/pii/S221478291500010X

166. Sainsbury K, Mullan B, Sharpe L. Dissemination of an Online Theory-Based Intervention to Improve Gluten-Free Diet Adherence in Coeliac Disease: the Relationship Between Acceptability, Effectiveness, and Attrition. Int J Behav Med 2015;22:356–364.

167. VanDeMark NR, Burrell NR, Lamendola WF, Hoich CA, Berg NP, Medina E. An exploratory study of engagement in a technology-supported substance abuse intervention. Subst Abuse Treat Prev Policy 2010;5(10):1–14.

168. Al-Asadi AM, Klein B, Meyer D. Pretreatment attrition and formal withdrawal during treatment and their predictors: an exploratory study of the anxiety online data. J Med Internet Res [Internet] Journal of Medical Internet Research; 2014 Jan 17 [cited 2016 Feb 5];16(6):e152. Available from: http://www.jmir.org/2014/6/e152/ PMID: 24938311

169. Habibović M, Cuijpers P, Alings M, van der Voort P, Theuns D, Bouwels L, et al. Attrition and adherence in a WEB-Based Distress Management Program for Implantable Cardioverter defibrillator Patients (WEBCARE): randomized controlled trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2014 Jan 28 [cited 2016 Feb 5];16(2):e52. Available from: http://www.jmir.org/2014/2/e52/ PMID: 24583632

170. Hebert EA, Vincent N, Lewycky S, Walsh K. Attrition and Adherence in the Online Treatment of Chronic Insomnia. Behav Sleep Med 2010;8(3):141–150.

171. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial Web-based weight loss program. J Med Internet Res [Internet] Journal of Medical Internet Research; 2010 Jan 14 [cited 2016 Jan 22];12(4):e69. Available from: http://www.jmir.org/2010/4/e69/ PMID: 21156470

172. Nicholas J, Proudfoot J, Parker G, Gillis I, Burckhardt R, Manicavasagar V, et al. The ins and outs of an online bipolar education program: a study of program attrition. J Med Internet Res [Internet] Journal of Medical Internet Research; 2010 Jan 19 [cited 2016 Feb 8];12(5):e57. Available from: http://www.jmir.org/2010/5/e57/ PMID: 21169169

173. Hong J-C, Chiu P-Y, Shih H-F, Lin P-S. Computer self-efficacy, competitive anxiety and flow state: Escaping from firing online game. Turkish Online J Educ Technol 2012;11(3):70–76.

174. Meischke H, Lozano P, Zhou C, Garrison MM, Christakis D. Engagement in “My Child’s Asthma”, an interactive web-based pediatric asthma management intervention. Int J Med Inform 2011;80(11):765–774.

175. Boyle EA, Connolly TM, Hainey T, Boyle JM. Engagement in digital entertainment games: A systematic review. Comput Human Behav [Internet] Elsevier Ltd; 2012;28(3):771–780. Available from: http://dx.doi.org/10.1016/j.chb.2011.11.020

176. Haines-Saah RJ, Kelly MT, Oliffe JL, Bottorff JL. Picture Me Smokefree: a qualitative study using social media and digital photography to engage young adults in tobacco reduction and cessation. J Med Internet Res [Internet] Journal of Medical Internet Research; 2015 Jan 26 [cited 2016 Jan 15];17(1):e27. Available from: http://www.jmir.org/2015/1/e27/ PMID: 25624064

177. Kim YH, Kim DJ, Wachter K. A study of mobile user engagement (MoEN): Engagement motivations, perceived value, satisfaction, and continued engagement intention. Decis Support Syst 2013;56:361–370.

178. Parks AC. A case for the advancement of the design and study of online positive psychological interventions. J Posit Psychol 2014;9(6):502–508.

179. Horsch C, Lancee J, Beun RJ, Neerincx MA, Brinkman W-P. Adherence to Technology-Mediated Insomnia Treatment: A Meta-Analysis, Interviews, and Focus Groups. J Med Internet Res [Internet] Journal of Medical Internet Research; 2015 Jan 4 [cited 2016 Feb 5];17(9):e214. Available from: http://www.jmir.org/2015/9/e214/ PMID: 26341671

180. Funk KL, Stevens VJ, Appel LJ, Bauck A, Brantley PJ, Champagne CM, et al. Associations of internet website use with weight change in a long-term weight loss maintenance program. J Med Internet Res [Internet] 2010 Jan [cited 2015 Nov 3];12(3):e29. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2956327&tool=pmcentrez&rendertype=abstract PMID: 20663751

181. Graham AL, Cha S, Cobb NK, Fang Y, Niaura RS, Mushro A. Impact of

seasonality on recruitment, retention, adherence, and outcomes in a web-based smoking cessation intervention: randomized controlled trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2013 Jan 7 [cited 2016 Feb 5];15(11):e249. Available from: http://www.jmir.org/2013/11/e249/ PMID: 24201304

182. Peels DA, Bolman C, Golsteijn RHJ, De Vries H, Mudde AN, van Stralen MM, et al. Differences in reach and attrition between Web-based and print-delivered tailored interventions among adults over 50 years of age: clustered randomized trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2012 Jan 17 [cited 2016 Feb 8];14(6):e179. Available from: http://www.jmir.org/2012/6/e179/ PMID: 23246790

183. Steinberg DM, Levine EL, Lane I, Askew S, Foley PB, Puleo E, et al. Adherence to self-monitoring via interactive voice response technology in an eHealth intervention targeting weight gain prevention among Black women: randomized controlled trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2014 Jan 29 [cited 2016 Feb 8];16(4):e114. Available from: http://www.jmir.org/2014/4/e114/ PMID: 24780934

184. Wanner M, Martin-Diener E, Bauer G, Braun-Fahrländer C, Martin BW. Comparison of trial participants and open access users of a web-based physical activity intervention regarding adherence, attrition, and repeated participation. J Med Internet Res [Internet] Journal of Medical Internet Research; 2010 Jan 10 [cited 2016 Feb 8];12(1):e3. Available from: http://www.jmir.org/2010/1/e3/ PMID: 20147006

185. Jahangiry L, Shojaeizadeh D, Montazeri A, Najafi M. Adherence and

Attrition in a Web-Based Lifestyle Intervention for People with Metabolic Syndrome. Iran J Public Health 2014;43(9):1248–1258.

186. Kuijpers W, Groen WG, Aaronson NK, van Harten WH. A systematic review of web-based interventions for patient empowerment and physical activity in chronic diseases: relevance for cancer survivors. J Med Internet Res [Internet] Journal of Medical Internet Research; 2013 Jan 20 [cited 2015 Jul 24];15(2):e37. Available from: http://www.jmir.org/2013/2/e37/ PMID: 23425685

187. Mahmassani HS, Chen RB, Huang Y, Williams D, Contractor N. Time to Play? Activity Engagement in Multiplayer Online Role-Playing Games. Transp Res Rec J Transp Res Board 2010;2157:129–137.

188. Ferguson MA, Henshaw H. Computer and Internet Interventions to Optimize Listening and Learning for People With Hearing Loss: Accessibility, Use, and Adherence. Am J Audiol 2015;24:338–343.

189. Weston A, Morrison L, Yardley L, Kleek M Van, Weal M. Measurements of engagement in mobile behavioural interventions? Digit Heal 2015. p. 1–8.

190. Donovan E, Mahapatra P Das, Green TC, Chiauzzi E, Mchugh K, Hemm A, et al. Efficacy of an online intervention to reduce alcohol-related risks among community college students. Addict Res Theory 2015;23(5):437–447.

191. Khadjesari Z, Murray E, Kalaitzaki E, White IR, McCambridge J,

Thompson SG, et al. Impact and costs of incentives to reduce attrition in online trials: two randomized controlled trials. J Med Internet Res [Internet] Journal of Medical Internet Research; 2011 Jan 2 [cited 2016 Feb 5];13(1):e26. Available from: http://www.jmir.org/2011/1/e26/ PMID: 21371988

192. An LC, Perry CL, Lein EB, Klatt C, Farley DM, Bliss RL, et al. Strategies for increasing adherence to an online smoking cessation intervention for college students. Nicotine Tob Res 2006;8(December):S7–S12. PMID: 17491165

193. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res [Internet] 2011 Jan [cited 2015 Mar 6];13(1):e2. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3221341&tool=pmcentrez&rendertype=abstract PMID: 21212045

194. Cairns P, Cox AL, Day M, Martin H, Perryman T. Who but not where: The effect of social play on immersion in digital games. Int J Hum Comput Stud [Internet] Elsevier; 2013;71:1069–1077. Available from: http://dx.doi.org/10.1016/j.ijhcs.2013.08.015

195. Morris RR, Schueller SM, Picard RW. Efficacy of a Web-based, crowdsourced peer-to-peer cognitive reappraisal platform for depression: randomized controlled trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2015 Jan 30 [cited 2015 Dec 17];17(3):e72.

Available from: http://www.jmir.org/2015/3/e72/ PMID: 25835472

196. Crutzen R, Cyr D, Larios H, Ruiter RAC, Vries NK De. Social Presence and Use of Internet-Delivered Interventions: A Multi-Method Approach. PLoS One 2013;8(2):e57067.

197. Ben-Zeev D, Kaiser SM, Krzos I. Remote “Hovering” with Individuals with Psychotic Disorders and Substance Use: Feasibility, Engagement, and Therapeutic Alliance with a Text-Messaging Mobile Interventionist. J Dual Diagn 2014;10(4):197–203.

198. Miller AS, Cafazzo JA, Seto E. A game plan: Gamification design principles in mHealth applications for chronic disease management. Health Informatics J 2014;1–10.

199. Brigham TJ. An Introduction to Gamification: Adding Game Elements for Engagement. Med Ref Serv Q [Internet] 2015;34(4):471–480. Available from: http://www.tandfonline.com.uml.idm.oclc.org.proxy2.lib.umanitoba.ca/doi/abs/10.1080/02763869.2015.1082385#.VnriLPkrJhE PMID: 26496401

200. Richardson CR, Buis LR, Janney AW, Goodrich DE, Sen A, Hess ML, et al. An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2010 Jan 17 [cited 2016 Jan 22];12(4):e71. Available from: http://www.jmir.org/2010/4/e71/ PMID: 21169160

201. Leslie E, Marshall AL, Owen N, Bauman A. Engagement and retention of participants in a physical activity website. Preventive 2005;40:54–59.

202. Irvine AB, Russell H, Manocchia M, Mino DE, Cox Glassen T, Morgan R, et al. Mobile-Web app to self-manage low back pain: randomized controlled trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2015 Jan 2 [cited 2016 Jan 5];17(1):e1. Available from: http://www.jmir.org/2015/1/e1/ PMID: 25565416

203. Lin H, Wu X. Intervention Strategies for Improving Patient Adherence to Follow-Up in the Era of Mobile Information Technology: A Systematic Review and Meta-Analysis. PLoS One 2014;9(8):e104266.

204. Kok G, Bockting C, Burger H, Smit F, Riper H. Mobile Cognitive Therapy: Adherence and acceptability of an online intervention in remitted recurrently depressed patients. Internet Interv [Internet] The Authors; 2014;1:65–73. Available from: http://dx.doi.org/10.1016/j.invent.2014.05.002

205. van den Berg MH, Ronday HK, Peeters AJ, Voogt-van der Harst EM, Munneke M, Breedveld FC, et al. Engagement and satisfaction with an Internet-based physical activity intervention in patients with rheumatoid arthritis. Rheumatology 2007;46(3):545–552. PMID: 17043050

206. Stark S, Snetselaar L, Piraino B, Stone A, Kim S, Hall B, et al. PDA self-monitoring adherence rates in two dialysis dietary intervention pilot studies: BalanceWise-HD and BalanceWise-PD. J Ren Nutr 2011;21(6):492–498.

207. Mohr DC, Duffecy J, Ho J, Kwasny M, Cai X, Burns MN, et al. A Randomized Controlled Trial Evaluating a Manualized TeleCoaching Protocol for Improving Adherence to a Web-Based Intervention for the Treatment of Depression. PLoS One 2013;8(8):e70086.

208. Klein M, Mogles N, Wissen A Van. Intelligent mobile support for therapy adherence and behavior change. J Biomed Inform [Internet] Elsevier Inc.; 2014;51:137–151. Available from: http://dx.doi.org/10.1016/j.jbi.2014.05.005

209. McCambridge J, Kalaitzaki E, White IR, Khadjesari Z, Murray E, Linke S, et al. Impact of length or relevance of questionnaires on attrition in online trials: randomized controlled trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2011 Jan 18 [cited 2016 Feb 5];13(4):e96. Available from: http://www.jmir.org/2011/4/e96/ PMID: 22100793

210. Helander E, Kaipainen K, Korhonen I, Wansink B. Factors related to sustained use of a free mobile app for dietary self-monitoring with photography and peer feedback: retrospective cohort study. J Med Internet Res [Internet] Journal of Medical Internet Research; 2014 Jan 15 [cited 2016 Feb 5];16(4):e109. Available from: http://www.jmir.org/2014/4/e109/ PMID: 24735567

211. Whiteside U, Lungu A, Richards J, Simon GE, Clingan S, Siler J, et al. Designing messaging to engage patients in an online suicide prevention intervention: survey results from patients with current suicidal ideation. J Med Internet Res [Internet] Journal of Medical Internet Research; 2014 Jan 7 [cited 2016 Feb 8];16(2):e42. Available from:

http://www.jmir.org/2014/2/e42/ PMID: 24509475

212. Jennings M. Theory and models for creating engaging and immersive e-commerce websites. Proc 2000 ACM SIGCPR Conf Comput Pers Res New York: ACM; 2000. p. 77–85.

213. Park N, Min K, Jin SA, Kang S. Effects of pre-game stories on feelings of presence and evaluation of computer games. Int J Hum Comput Stud [Internet] Elsevier; 2010;68:822–833. Available from: http://dx.doi.org/10.1016/j.ijhcs.2010.07.002

214. Hwang M-Y, Hong J-C, Hao Y-W, Jong J-T. Elders' Usability, Dependability, and Flow Experiences on Embodied Interactive Video Games. Educ Gerontol 2011;37(8):715–731.

215. Chapman P, Selvarajah S, Webster J. Engagement in Multimedia Training Systems. Proc 32nd Hawaii Int Conf Syst Sci Washington, DC: IEEE; 1999. p. 1–9.

216. Liu S, Liao H, Pratt JA. Impact of media richness and flow on e-learning technology acceptance. Comput Educ [Internet] Elsevier Ltd; 2009;52:599–607. Available from: http://dx.doi.org/10.1016/j.compedu.2008.11.002

217. Miller AS, Cafazzo JA, Seto E. A game plan: Gamification design principles in mHealth applications for chronic disease management. Health Informatics J [Internet] 2014 Jul 1 [cited 2015 Oct 1]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/24986104 PMID: 24986104

218. Lieberman DZ. Effects of a Personified Guide on Adherence to an Online Program for Alcohol Abusers. Cyberpsychology Behav 2006;9(5):603–607.

219. Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, et al. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. Heal Psychol [Internet] 2004 Sep [cited 2015 Apr 25];23(5):443–451. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15367063 PMID: 15367063

220. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. J Public Health Dent 2011;71:S52–S63. PMID: 21499543

221. Ubhi HK, Michie S, Kotz D, Wong WC, West R. A mobile app to aid smoking cessation: Preliminary evaluation of SmokeFree28. J Med Internet Res 2015;17(1):e17. PMID: 25596170

222. Chinn PL, Kramer MK. Theory and nursing: a systematic approach. St. Louis: Mosby-Year Book; 1991.

223. Fiannaca A, La Rosa M, Rizzo R, Urso A, Gaglio S. An ontology design methodology for Knowledge-Based systems with application to bioinformatics. Comput Intell Bioinforma Comput Biol (CIBCB), 2012 IEEE Symp 2012. p. 85–91.

224. Weber R. Evaluating and Developing Theories in the Information Systems

Discipline. J Assoc Inf Syst 2012;13(1):1–30.

225. O'Brien HL, Toms EG. The Development and Evaluation of a Survey to Measure User Engagement. J Am Soc Inf Sci Technol 2010;61(1):50–69.

226. Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavorial medicine. Ann Behav Med 1994;16(3):199–202.

227. Haukkala A, Uutela A, Vartiainen E, Mcalister A, Knekt P. Depression and smoking cessation: The role of motivation and self-efficacy. Addict Behav. 2000.

228. Linde JA, Jeffery RW, Levy RL, Sherwood NE, Utter J, Pronk NP, et al. Binge eating disorder, weight control self-efficacy, and depression in overweight men and women. Int J Obes [Internet] Nature Publishing Group; 2004 Mar 13 [cited 2016 Jun 10];28(3):418–425. Available from: http://www.nature.com/doifinder/10.1038/sj.ijo.0802570

229. Public Health England. One You [Internet]. 2016 [cited 2016 Oct 20]. Available from: https://www.nhs.uk/oneyou/apps

230. National Instítues of Health. National Library of Medicine Mobile [Internet]. 2016 [cited 2016 Oct 20]. Available from: https://www.nlm.nih.gov/mobile/

231. The Nielsen Company. The Mobile Media Report - State Of The Media - Q3 2011. 2011.

232. Zhang MWB, Ward J, Ying JJB, Pan F, Ho RCM. The alcohol tracker application: an initial evaluation of user preferences. BMJ Innov

2016;2(1):8–13. PMID: 27019744

233. Garnett C, Crane D, Michie S, West R, Brown J. Evaluating the effectiveness of a smartphone app to reduce excessive alcohol consumption: protocol for a factorial randomised control trial. BMC Public Health [Internet] BMC Public Health; 2016;16(1):536. Available from: http://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-016-3140-8 PMID: 27392430

234. Ubhi HK, Michie S, Kotz D, Wong WC, West R. A Mobile App to Aid Smoking Cessation: Preliminary Evaluation of SmokeFree28. J Med Internet Res [Internet] 2015 Jan 16 [cited 2015 Mar 11];17(1):e17. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4319069&tool=pmcentrez&rendertype=abstract PMID: 25596170

235. Zhang MWB, Ho CSH, Cheok CCS, Ho RCM. Smartphone apps in mental healthcare: the state of the art and potential developments. BJPscyh Adv [Internet] 2015;21(5):354–358. Available from: http://apt.rcpsych.org/cgi/doi/10.1192/apt.bp.114.013789

236. Lindgaard G, Fernandes G, Dudek C, Brown J. Attention web designers: You have 50 milliseconds to make a good first impression! Behav Inf Technol 2006;25(2):115–126. PMID: 19277458

237. Norman DA. Emotional design: why we love (or hate) everyday things. New York: Basic Books; 2004.

238. Ayalew R. Consumer behaviour in Apple's App Store. Uppsala University; 2011.

239. Chang T-R, Kaasinen E, Kaipainen K. What Influences Users' Decisions to Take Apps into Use? A Framework for Evaluating Persuasive and Engaging Design in Mobile Apps for Well-Being. Proc 11th Int Conf Mob Ubiquitous Multimed Ulm, Germany: ACM; 2012. p. 2.

240. Heffner JL, Vilardaga R, Mercer LD, Kientz J a., Bricker JB. Feature-level analysis of a novel smartphone application for smoking cessation. Am J Drug Alcohol Abuse [Internet] Informa Healthcare USA, Inc; 2015;41(1):68–73. Available from: http://informahealthcare.com/doi/abs/10.3109/00952990.2014.977486

241. Zhang MWB, Fang P, Ho RCM. Global outreach and user preferences of a smartphone application developed for drinkers. Technol Heal Care 2016;24(4):495–501. PMID: 26890229

242. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013;15(4):1–17. PMID: 23598614

243. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Heal Care 2007;19(6):349–357. PMID: 17872937

244. Ericsson AK, Simon HA. How to study thinking in everyday life:

Contrasting think-aloud protocols with descriptions and explanations of thinking. Mind, Cult Act 1998;5(3):178–186.

245. De Ridder D. Nudging for beginners: A shortlist of issues in urgent need of research. Eur Heal Psychol 2014;16(1):2–6.

246. Maguire M. Methods to support human-centred design. Int J Hum Comput Stud [Internet] 2001;55(4):587–634. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1071581901905038 PMID: 21871822

247. Yardley L, Morrison LG, Andreou P, Joseph J, Little P. Understanding reactions to an internet-delivered health-care intervention: accommodating user preferences for information provision. BMC Med Informatics Decis [Internet] 2010;10:52. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2946266&tool=pmcentrez&rendertype=abstract PMID: 20849599

248. Yardley L, Morrison L, Bradbury K, Muller I. The Person-Based Approach to Intervention Development: Application to Digital Health-Related Behavior Change Interventions. J Med Internet Res 2015;17(1):e301.

249. Schwandt TA. Constructivist, Interpretivist Approaches to Human Inquiry. In: Denzin NK, Lincoln YS, editors. Handb Qual Res London: SAGE Publications; 1994. p. 118–137.

250. Ponterotto JG. Qualitative Research in Counseling Psychology: A Primer on Research Paradigms and Philosophy of Science. J Couns Psychol

2005;52(2):126–136.

251. Bradley KA, Debenedetti AF, Volk RJ, Williams EC, Frank D, Kivlahan DR. AUDIT-C as a brief screen for alcohol misuse in primary care. Alcohol Clin Exp Res 2007;31(7):1208–1217. PMID: 17451397

252. Rosson MB, Carroll JM. Usability Engineering: Scenario-Based Development of Human-Computer Interaction. Morgan Kaufmann; 2002.

253. Brown J, Michie S, Raupach T, West R. Prevalence and characteristics of smokers interested in internet-based smoking cessation interventions: cross-sectional findings from a national household survey. J Med Internet Res [Internet] 2013 Jan [cited 2015 Aug 4];15(3):e50. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3636298&tool=pmcentrez&rendertype=abstract PMID: 23506944

254. Biernacki P, Waldorf D. Snowball sampling: Problems and Techniques of Chain Referral Sampling. Sociol Methods Res [Internet] 1981;10(2):141–163. Available from: http://www.uk.sagepub.com/gray3e/study/chapter9/Encyclopaedia entries/Snowball_Sampling.pdf

255. Morse JM. Theoretical Saturation. In: Lewis-Beck MS, Bryman A, Futing Liao T, editors. SAGE Encycl Soc Sci Res Methods Thousand Oaks: SAGE Publications, Inc.; 2004.

256. Office for National Statistics. Guidance and Methodology: Ethnic Group, National Identity and Religion [Internet]. 2016 [cited 2016 Mar 25].

Available from:
http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/measuring-equality/equality/ethnic-nat-identity-religion/ethnic-group/index.html

257. Chandola T, Jenkinson C. The new UK National Statistics Socio-Economic Classification (NS-SEC); investigating social class differences in self-reported health status. J Public Health Med 2000;22(2):182–190. PMID: 10912557

258. Borland R, Yong HH, O'Connor RJ, Hyland A, Thompson ME. The reliability and predictive validity of the heaviness of smoking index and its two components: Findings from the International Tobacco Control Four Country Study. Nicotine Tob Res 2010;12:45–50. PMID: 20889480

259. Heatherton TF, Kozlowski LT, Frecker RC, Rickert W, Robinson J. Measuring the heaviness of smoking: using self-reported time to the first cigarette of the day and number of cigarettes smoked per day. Br J Addict 1989;84(7):791–799. PMID: 2758152

260. Bush KR, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT Alcohol Consumption Questions (AUDIT-C). Arch Intern Med 1998;158:1789–1795.

261. Frank D, DeBenedetti AF, Volk RJ, Williams EC, Kivlahan DR, Bradley KA. Effectiveness of the AUDIT-C as a screening test for alcohol misuse in three race/ethnic groups. J Gen Intern Med 2008;23(6):781–787. PMID: 18421511

262. Kotz D, Brown J, West R. Predictive validity of the Motivation To Stop Scale (MTSS): A single-item measure of motivation to stop smoking. Drug Alcohol Depend [Internet] Elsevier Ireland Ltd; 2013;128(1–2):15–19. Available from: http://dx.doi.org/10.1016/j.drugalcdep.2012.07.012

263. Rollnick S, Heather N, Gold R, Hall W. Development of a short “readiness to change” questionnaire for use in brief, opportunistic interventions among excessive drinkers. Br J Addict 1992;87:743–754.

264. West R. Time for a change: putting the Transtheoretical (Stages of Change) Model to rest. Addiction 2005;1036–1039.

265. de Vocht F, Brown J, Beard E, Angus C, Brennan A, Michie S, et al. Temporal patterns of alcohol consumption and attempts to reduce alcohol intake in England. BMC Public Health [Internet] BMC Public Health; 2016;16(1):917. Available from: http://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-016-3542-7

266. Qualtrics. Qualtrics Survey Software [Internet]. Provo, Utah; 2016. Available from: http://www.qualtrics.com/

267. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(May 2015):77–101. PMID: 223135521

268. QSR International. NVivo 10. 2012.

269. Lincoln YS, Guba EG. Naturalistic Inquiry. SAGE Publications; 1985.

270. Mays N, Pope C. Assessing quality in qualitative research. Br Med J 2000;320:50–52. PMID: 10617534

271. Graham HM. Smoking, Stigma and Social Class. J Soc Polit 2012;41(1):83–99.

272. Room R. Stigma, social inequality and alcohol and drug use. Drug Alcohol Rev 2005;24(2):143–155. PMID: 16076584

273. Department of Health. Research Governance Framework for Health and Social Care (Second Edition). 2005. PMID: 11892625

274. Spruijt-Metz D, Nilsen W. Dynamic Models of Behavior for Just-in-Time Adaptive Interventions. IEEE Pervasive Comput 2014;3:13–17.

275. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011;1:53–71.

276. Naughton F, Hopewell S, Lathia N, Schalbroeck R, Brown C, Mascolo C, et al. A Context-Sensing Mobile Phone App (Q Sense) for Smoking Cessation: A Mixed-Methods Study. JMIR mHealth uHealth [Internet] JMIR mHealth and uHealth; 2016 Sep 16 [cited 2016 Sep 16];4(3):e106. Available from: http://mhealth.jmir.org/2016/3/e106/

277. Carpenter MJ, Sterba KR, Boatright AS, West R. “Closet” quit attempts: Prevalence, correlates and association with outcome. Addiction 2011;106(12):2214–2220. PMID: 21672072

278. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. JMIR mHealth uHealth [Internet] 2015;3(1):e27. Available from: http://mhealth.jmir.org/2015/1/e27/ PMID: 25760773

279. Ubhi HK, Michie S, Phil D, Kotz D, Van Schayck OCP, Selladurai A, et al. Characterising smoking cessation smartphone applications in terms of behaviour change techniques, engagement and ease-of-use features. Transl Behav Med 2015;1–8.

280. Branch JL. Investigating the Information-Seeking Processes of Adolescents. Libr Inf Sci Res 2000;22(4):371–392.

281. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med [Internet] Springer US; 2017 Dec 13 [cited 2016 Dec 19];7:254–267. Available from: http://link.springer.com/10.1007/s13142-016-0453-1

282. Zaichkowsky JL. Measuring the Involvement Construct. J Consum Res 1985;12(3):341–352.

283. McClure JB, Shortreed SM, Bogart A, Derry H, Riggs K, St John J, et al. The effect of program design on engagement with an internet-based smoking intervention: randomized factorial trial. J Med Internet Res [Internet] Journal of Medical Internet Research; 2013 Jan 25 [cited 2016 Feb 5];15(3):e69. Available from: http://www.jmir.org/2013/3/e69/ PMID: 23529377

284. O'Brien HL, Toms EG. Examining the generalizability of the User Engagement Scale (UES) in exploratory search. Inf Process Manag [Internet] Elsevier Ltd; 2013;49(5):1092–1107. Available from: http://dx.doi.org/10.1016/j.ipm.2012.08.005

285. Zaichkowsky JL. The Personal Involvement Inventory: Reduction, Revision, and Application to Advertising. J Advert 1994;23(4):59–70. PMID: 9502063976

286. MacKenzie S, Podsakoff P, Podsakoff N. Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques. MIS Q [Internet] 2011;35(2):293–334. Available from: http://www.jstor.org/stable/10.2307/23044045 PMID: 60461934

287. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak BMC Medical Informatics and Decision Making; 2017;17(25):1–14.

288. Milward J, Drummond C, Fincham-Campbell S, Deluca P. What makes online substance-use interventions engaging? A systematic review and narrative synthesis. Digit Heal [Internet] 2018;4:1–25. Available from: http://journals.sagepub.com/doi/10.1177/2055207617743354

289. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res [Internet] Journal of Medical Internet

Research; 2011 Jan 6 [cited 2016 Jan 9];13(1):e2. Available from: http://www.jmir.org/2011/1/e2/ PMID: 21212045

290. Chaplin WF, John OP, Goldberg LR. Conceptions of States and Traits: Dimensional Attributes With Ideals as Prototypes. J Pers Soc Psychol 1988;54(4):541–557. PMID: 3367279

291. Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: The development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative da. Health Technol Assess (Rockv) [Internet] 2015;19(99):1–187. Available from: http://dx.doi.org/10.3310/hta19990 PMID: 26616119

292. Michie S, Abraham C, Whittington C, Mcateer J. Effective Techniques in Healthy Eating and Physical Activity Interventions: A Meta-Regression. Heal Psychol 2009;28(6):690–701.

293. Michie S, Whittington C, Hamoudi Z, Zarnani F, Tober G, West R. Identification of behaviour change techniques to reduce excessive alcohol consumption. Addiction 2012;107(8):1431–1440. PMID: 22340523

294. Posner MI, Petersen SE. The Attention System of the Human Brain. Annu Rev Neurosci [Internet] 1990;13:25–42. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2183676 PMID: 2183676

295. Yiend J. The effects of emotion on attention: A review of attentional

processing of emotional information. Cogn Emot 2010;24(1):3–47.

296. Tomkins SS. Affect Imagery Consciousness: The Complete Edition [Internet]. New York: Springer Publishing Company; 2008. Available from: http://books.google.com/books?id=l4GC6rqlmD0C PMID: 15112580

297. Silvia PJ. Interest - The Curious Emotion. Curr Dir Psychol Sci 2008;17(1):57–60.

298. Hinkin TR. A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires. Organ Res Methods 1998;1(1):104–121.

299. Weng LJ. Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. Educ Psychol Meas 2004;64(6):956–972.

300. MacKenzie SB, Podsakoff PM, Fetter R. Organizational citizenship behavior and objective productivity as determinants of managerial evaluations of salespersons’ performance. Organ Behav Hum Decis Process 1991;50:123–150. PMID: 5027403

301. Hinkin TR, Schriesheim CA. Development and Application of New Scales to Measure the French and Raven (1959) Bases of Social Power. J Appl Psychol 1989;74(4):561–567.

302. Google Analytics. How a web session is defined in Analytics [Internet]. 2017 [cited 2018 Feb 6]. Available from: https://support.google.com/analytics/answer/2731565

303. Fan W, Yan Z. Factors affecting response rates of the web survey: A systematic review. Comput Human Behav [Internet] Elsevier Ltd; 2010;26(2):132–139. Available from: http://dx.doi.org/10.1016/j.chb.2009.10.015

304. IBM Corp. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corporation; 2012.

305. Costello AB, Osborne JW. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. Pract Assessment, Res Eval [Internet] 2005;10(7):1–9. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.110.9154&rep=rep1&type=pdf PMID: 12345678

306. Morrison L, Geraghty AWA, Lloyd S, Goodman N, Michaelides D, Hargood C, et al. Comparing usage of a web- and smartphone app-delivered stress management intervention: An observational study. Internet Interv.

307. Kaiser HF. An index of factorial simplicity. Psychometrika 1974;39(1):31–36.

308. Lakens D. Observed power, and what to do if your editor asks for post-hoc power analyses [Internet]. 20% Stat. 2014 [cited 2018 Sep 9]. Available from: http://daniellakens.blogspot.com/2014/12/observed-power-and-what-to-do-if-your.html

309. Willis GB. Cognitive interviewing: A tool for improving questionnaire

design. SAGE Publications, Inc.; 2004.

310. Radtke T, Ostergaard M, Cooke R, Scholz U. Web-based alcohol intervention: Study of systematic attrition of heavy drinkers. J Med Internet Res 2017;19(6):1–12. PMID: 28659251

311. Peer E, Brandimarte L, Samat S, Acquisti A. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. J Exp Soc Psychol [Internet] Elsevier Inc.; 2017;70:153–163. Available from: http://dx.doi.org/10.1016/j.jesp.2017.01.006

312. Hummel K, Brown J, Willemsen MC, West R, Kotz D. External validation of the motivation to stop scale (MTSS): Findings from the international tobacco control (ITC) Netherlands survey. Eur J Public Health 2016;27(1):129–134. PMID: 27452891

313. Mundfrom DJ, Shaw DG, Ke TL. Minimum Sample Size Recommendations for Conducting Factor Analyses. Int J Test [Internet] 2005;5(2):159–168. Available from: http://dx.doi.org/10.1207/s15327574ijt0502

314. Palan S, Schitter C. Prolific.ac - A subject pool for online experiments. J Behav Exp Financ [Internet] Elsevier B.V.; 2018;17:22–27. Available from: https://doi.org/10.1016/j.jbef.2017.12.004

315. Paolacci G, Chandler J, Ipeirotis P. Running experiments on Amazon Mechanical Turk. Judgm Decis Mak [Internet] 2010;5(5):411–419. Available from: http://repub.eur.nl/pub/31983 PMID: 1011330798

316. Gustafson DH, Shaw BR, Isham A, Baker T, Boyle MG, Levy M. Explicating an evidence-based, theoretically informed, mobile technology-based system to improve outcomes for people in recovery for alcohol dependence. Subst Use Misuse 2011;46(1):96–111. PMID: 21190410

317. Mellentin AI, Stenager E, Nielsen B, Nielsen AS, Yu F. A Smarter Pathway for Delivering Cue Exposure Therapy? The Design and Development of a Smartphone App Targeting Alcohol Use Disorder. JMIR mHealth uHealth [Internet] 2017;5(1):e5. Available from: http://mhealth.jmir.org/2017/1/e5/ PMID: 28137701

318. Hamilton FL, Hornby J, Sheringham J, Linke S, Ashton C, Moore K, et al. DIAMOND (DIgital Alcohol Management ON Demand): a mixed methods feasibility RCT and embedded process evaluation of a digital health intervention to reduce hazardous and harmful alcohol use. Pilot Feasibility Stud Pilot and Feasibility Studies; 2017;34(3):1–12.

319. Kitzinger J. The methodology of focus groups: the importance of interaction between research participants. Sociol Health Illn 1994;16(1):103–121. PMID: 9

320. Morgan DL. What Do You Get from Focus Groups? Focus Gr Guideb Thousand Oaks: SAGE Publications, Inc.; 1998. p. 55–64.

321. Krosnick J. Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. Appl Cogn Psychol 1991;5:213–236.

322. Oppenheimer DM, Meyvis T, Davidenko N. Instructional manipulation

checks: Detecting satisficing to increase statistical power. J Exp Soc Psychol [Internet] Elsevier Inc.; 2009;45:867–872. Available from: http://dx.doi.org/10.1016/j.jesp.2009.03.009

323. Deci EL, Ryan RM. Self-Determination Theory: A Macrotheory of Human Motivation, Development, and Health. Can Psychol 2008;49(3):182–185.

324. Deci EL, Ryan RM, Koestner R. A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. Psychol Bull 1999;125(6):627–668.

325. Consolvo S, McDonald DW, Landay JA. Theory-Driven Design Strategies for Technologies that Support Behavior Change in Everyday Life. CHI ’09 - Creat Thought Self-Improvement 2009. p. 405–414.

326. Milward J, Khadjesari Z, Fincham-Campbell S, Deluca P, Watson R, Drummond C. User Preferences for Content, Features, and Style for an App to Reduce Harmful Drinking in Young Adults: Analysis of User Feedback in App Stores and Focus Group Interviews. JMIR mHealth uHealth 2016;4(2):e47.

327. Cooper A, Reimann R, Cronin D, Noessel C. About Face: The Essentials of Interaction Design. 4th ed. Wiley-Blackwell; 2014.

328. Yardley L, Bishop FL, Beyer N, Hauer K, Kempen GI, Piot-Ziegler C, et al. Older people’s views of falls-prevention interventions in six European countries. Gerontologist 2006;46(5):650–660.

329. Wilsnack RW, Vogeltanz ND, Wilsnack SC, Harris TR. Gender differences in alcohol consumption and adverse drinking consequences: cross-cultural patterns. Addiction 2000;95(2):251–265.

330. Nolen-Hoeksema S, Hilt L. Possible contributors to the gender differences in alcohol use and problems. J Gen Psychol 2006;133(4):357–374. PMID: 17128956

331. Johnston DW, Johnston M. Useful theories should apply to individuals. Br J Health Psychol 2013;18(3):469–473. PMID: 23724956

332. Molenaar PCM. A Manifesto on Psychology as Idiographic Science: Bringing the Person Back Into Scientific Psychology, This Time Forever. Measurement 2004;2(4):201–218.

333. Linke S, Murray E, Butler C, Wallace P. Internet-based interactive health intervention for the promotion of sensible drinking: Patterns of use and potential impact on members of the general public. J Med Internet Res 2007;9(2):1–12. PMID: 17513281

334. Garnett C, Perski O, Tombor I, West R, Michie S, Brown J. Predictors of engagement, response to follow-up and extent of alcohol reduction in users of a smartphone app, Drink Less. JMIR mHealth uHealth 2018;

335. Chiauzzi E, Green TC, Lord S, Thum C, Goldstein M. My student body: A high-risk drinking prevention website for college students. J Am Coll Heal 2005;53(6):263–274. PMID: 15900990

336. Perski O, Baretta D, Blandford A, West R, Michie S. Engagement features judged by excessive drinkers as most important to include in smartphone applications for alcohol reduction: A mixed methods study. Digit Heal.

337. Sedgwick P. What is an “n-of-1” trial? Br Med J 2014;348:1–2. PMID: 2014260588

338. Nisbett RE, Wilson TD. Telling more than we can know: Verbal reports on mental processes. Psychol Rev 1977;84(3):231–259. PMID: 17882490

339. Smyth JM, Stone AA. Ecological Momentary Assessment Research in Behavioral medicine. J Happiness Stud [Internet] 2003;4:35–52. Available from: http://www.springerlink.com/openurl.asp?id=doi:10.1023/A:1023657221954

340. Inauen J, Shrout PE, Bolger N, Stadler G, Scholz U. Mind the Gap? An Intensive Longitudinal Study of Between-Person and Within-Person Intention-Behavior Relations. Ann Behav Med [Internet] Annals of Behavioral Medicine; 2016;50(4):516–522. Available from: http://dx.doi.org/10.1007/s12160-016-9776-x

341. Duan N, Kravitz RL, Schmid CH. Single-patient (n-of-1) trials: A pragmatic clinical decision methodology for patient-centered comparative effectiveness research. J Clin Epidemiol [Internet] Elsevier Inc; 2013;66:S21–S28. Available from: http://dx.doi.org/10.1016/j.jclinepi.2013.04.006 PMID: 23849149

342. Kwasnicka D, Dombrowski SU, White M, Sniehotta FF. N-of-1 study of weight loss maintenance assessing predictors of physical activity, adherence to weight loss plan and weight change. Psychol Heal [Internet] Routledge; 2017;32(6):686–708. Available from: http://dx.doi.org/10.1080/08870446.2017.1293057 PMID: 28323457

343. Gilks W, Richardson S, Spiegelhalter D. Markov chain Monte Carlo in practice. CRC press; 1995.

344. Quinn F, Johnston M, Johnston DW. Testing an integrated behavioural and biomedical model of disability in N-of-1 studies with chronic pain. Psychol Heal 2013;28(12):1391–1406. PMID: 23863041

345. Hobbs N, Dixon D, Johnston M, Howie K. Can the theory of planned behaviour predict the physical activity behaviour of individuals? Psychol Health [Internet] 2013;28(3):234–249. Available from: http://www.tandfonline.com/doi/abs/10.1080/08870446.2012.716838 PMID: 22943555

346. Crane D, Garnett C, Brown J, West R, Michie S. Factors influencing usability of a smartphone app to reduce excessive alcohol consumption: Think aloud and interview studies. Front Public Heal 2017;5:1–19.

347. Shadish WR, Zuur AF, Sullivan KJ. Using generalized additive (mixed) models to analyze single case designs. J Sch Psychol [Internet] Society for the Study of School Psychology; 2014;52(2):149–178. Available from: http://dx.doi.org/10.1016/j.jsp.2013.11.004 PMID: 24606973

348. Akaike H. A New Look at the Statistical Model Identification. IEEE Trans Automat Contr 1974;19(6):716–723. PMID: 1100705

349. Vervloet M, Linn AJ, van Weert JCM, de Bakker DH, Bouvy ML, van Dijk L. The effectiveness of interventions using electronic reminders to improve adherence to chronic medication: A systematic review of the literature. J Am Med Informatics Assoc 2012;19:696–704.

350. Schepers J, Wetzels M. A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects. Inf Manag 2007;44(1):90–103. PMID: 25246403

351. Stawarz K, Rodríguez MD, Cox AL, Blandford A. Understanding the use of contextual cues: design implications for medication adherence technologies that support remembering. Digit Heal [Internet] 2016;2:1–18. Available from: http://journals.sagepub.com/doi/10.1177/2055207616678707 PMID: 29942574

352. Sieverink F, Kelders SM, Gemert-Pijnen V. Clarifying the concept of adherence to ehealth technology: Systematic review on when usage becomes adherence. J Med Internet Res 2017;19(12):e402. PMID: 29212630

353. Petrides K V., Furnham A. Trait emotional intelligence: Psychometric investigation with reference to established trait taxonomies. Eur J Pers 2001;15(6):425–448.

354. Blandford A, Gibbs J, Newhouse N, Perski O, Singh A, Murray E. Seven lessons for interdisciplinary research on interactive digital health interventions. Digit Heal [Internet] 2018;4:1–13. Available from: http://journals.sagepub.com/doi/10.1177/2055207618770325 PMID: 29942629

355. Cheong Y, Yong H, Borland R. Does how you quit affect success? A comparison between abrupt and gradual methods using data from the International Tobacco Control Policy Evaluation Study. Nicotine Tob Res [Internet] 2007;9(8):801–810. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17654293 PMID: 17654293

356. Bremer V, Becker D, Kolovos S, Funk B, van Breda W, Hoogendoorn M, et al. Predicting Therapy Success and Costs for Personalized Treatment Recommendations Using Baseline Characteristics: Data-Driven Analysis. J Med Internet Res [Internet] 2018;20(8):e10275. Available from: http://www.jmir.org/2018/8/e10275/ PMID: 30131318

357. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007;133(4):673–693. PMID: 17592961

358. Michie S, West R, Campbell R, Brown J, Gainforth H. ABC of Behaviour Change Theories. London: Silverback Publishing; 2014.

359. Hekler EB, Michie S, Pavel M, Rivera DE, Collins LM, Jimison HB, et al. Advancing Models and Theories for Digital Behavior Change Interventions. Am J Prev Med [Internet] Elsevier; 2016;51(5):825–832.

Available from: http://dx.doi.org/10.1016/j.amepre.2016.06.013 PMID: 27745682

Appendix 1 – Electronic search strategy (Study 1)

1. “user engagement”.ti,ab,sh.

2. engag*.ti,sh.

3. immersion.ti,sh.

4. flow.ti,sh.

5. involvement.ti,sh.

6. presence.ti,sh.

7. adherence.ti,sh.

8. attrition.ti,sh.

9. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8

10. digital.ti,sh.

11. web*.ti,sh.

12. computer.ti,sh.

13. online.ti,sh.

14. technology.ti,sh.

15. mobile.ti,sh.

16. smartphone.ti,sh.

17. 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16

18. “behavior?r change”.ti,ab,sh.

19. intervention.ti,ab,sh.

20. game*.ti,ab,sh.

21. multimedia.ti,ab,sh.

22. 18 OR 19 OR 20 OR 21

23. 9 AND 17 AND 22

Appendix 2 – Characteristics of included studies (Study 1)

Authors (Year)	Country	Study aim	Population	Technology	Programme length	Participant characteristics	Study design	Data collection method
Al-Asadi et al. (2014)	Australia	To identify predictors of pre-treatment attrition and formal withdrawal from the Anxiety Online program.	Anxiety	Website	12 weeks	N = 3,880; Mean age (SD) = 36.4 (12.1); % Female = 68.3	Cohort	Survey
An et al. (2006)	US	To identify rates of participation in the RealU intervention.	Smokers	Website	20 weeks	N = 257; Mean age (SD) = 20.1 (1.6); % Female = 70	Cohort	Survey
Arden-Close et al. (2015)	UK	To examine patterns of web usage amongst obese primary care patients within the POWeR intervention.	Obese individuals	Website	12 weeks	N= 132; Mean age (SD) = 51.6 (13.0); % Female = 66	Secondary analysis of RCT data, intervention arm	Website logs
Bellg et al. (2004)	UK	To conceptualise treatment fidelity and to offer recommendations for how to incorporate fidelity measures into intervention research.	N/A	N/A	N/A	N/A	Review, narrative synthesis	N/A
Ben-Zeev et al. (2014)	US	To assess the usability of and engagement with a mobile phone intervention.	Serious mental illness/substance abuse	Mobile phone	12 weeks	N = 17; Mean age (SD) = 40.5 (11.6); % Female = 41	Pre- posttest	Text-message log
Bianchi-Berthouze et al. (2007)	UK	To understand video game engagement based on body movements.	Healthy adults	Digital game	N/A	N = 14; Mean age (SD) = 25.0 (4.4)	Experimental, between-subjects design	Exoskeleton to measure upper body joint movement and video camera

Borrelli (2011)	US	To discuss the assessment, monitoring, and enhancement of treatment fidelity in public health trials.	N/A	N/A	N/A	N/A	Review, narrative synthesis	N/A
Bossen et al. (2013)	The Netherlands	To explore patient and study characteristics that facilitate or hinder usage of a Web-based physical activity intervention.	Patients diagnosed with hip and/or knee osteoarthritis	Website	9 weeks	N = 199; Mean age (SD) = 60.0 (6.3); % Female = 63	Mixed methods including secondary analysis of RCT data, intervention arm	Website logs and face-to-face interviews
Bouvier et al. (2014)	France	To gain a better understanding of what it means to be engaged and how to decide whether a behaviour reflects engagement or not.	Healthy adults	Digital game	N/A	N/A	Review, narrative synthesis	N/A
Boyle et al. (2012)	UK	To explore the diverse aspects of engagement and to develop a coherent understanding of engagement in computer games.	N/A	Digital game	N/A	N/A	Systematic review, narrative synthesis	N/A
Brigham (2015)	US	To explain the term ‘gamification’ and its current use.	N/A	N/A	N/A	N/A	Review, narrative synthesis	N/A
Brouwer et al. (2011)	The Netherlands	To identify methods that promote better exposure to internet interventions.	Primary prevention of physical chronic disease	Internet-delivered interventions	N/A	N/A	Systematic review, narrative synthesis	N/A
Brown & Cairns (2004)	UK	To develop a grounded theory of immersion.	Healthy adults	Digital game	N/A	N = 7; % Female = 43	Qualitative	Face-to-face interviews

Burns & Fairclough (2015)	UK	To quantify the degree of immersion in a digital world.	Healthy adults	Digital game	N/A	N = 20; Mean age (SD) = 23.7 (4.2); % Female = 35	Experimental, mixed design	Event-related potentials to task-irrelevant stimuli
Cairns et al. (2013)	UK	To explore how social play influences the immersive experience of digital gameplay.	Healthy adults	Digital game	N/A	N = 24; % Female = 42	Experimental, within-subjects design	Questionnaires
Calleja (2007)	Australia	To develop a conceptual model for understanding game involvement.	N/A	Digital game	N/A	N/A	Review, narrative synthesis	N/A
Carter et al. (2013)	UK	To compare the acceptability of a self-monitoring weight management intervention delivered by a smartphone app with that of a website.	Overweight individuals	Smartphone app, website	6 weeks	N = 128; Mean age (SD) = 41.2 (8.5); % Female = 77	RCT	Usage data
Chapman, Selvarajah, & Webster (1999)	Canada	To examine engagement in two types of multimedia training systems.	Healthy adults	Interactive software	N/A	N = 72; % Female = 69	Experimental, between-subjects design	Questionnaires
Chen et al. (2015)	US	To explore the nature of engagement with an online workshop for cancer survivors.	Cancer survivors	Web-based	8 weeks	N = 127; Mean age (range) = 52 (26-81); % Female = 82	Secondary analysis of RCT data, intervention arm	Usage data

Chiang et al. (2011)	Taiwan	To explore online game players' flow experiences.	Healthy adults	Digital game	N/A	N = 30; % Female = 63	Experimental, within-subjects design	Questionnaires
Chou et al. (2014)	Taiwan	To explore design factors that increase flow experience in mobile games.	Healthy adults	Mobile phone	N/A	N = 234	Qualitative	Focus groups
Christensen et al. (2009)	Australia	To review rates of adherence to internet interventions for anxiety and depression.	Anxiety, depression	Internet-based interventions	N/A	N/A	Systematic review, narrative synthesis	N/A
Chung & Gardner (2012)	Australia	To assess the effect of different kinds of induced interruptions on players' presence in a virtual reality theatre.	Healthy adults	Digital game	N/A	N = 36; Mean age (SD) = 22.4 (2.9)	Experimental, mixed design	Questionnaires
Couper et al. (2012)	US	To explore the qualities of engagement in an online intervention designed to promote fruit and vegetable consumption.	Healthy adults	Website	4 months	N = 2513; Mean age = 46.3; % Female = 69	RCT	Usage data
Crutzen et al. (2012)	The Netherlands	To assess whether user control increases website use.	Healthy adults	Website	N/A	N = 668; Mean age (SD) = 49.0 (16.0); % Female = 49.7	Experimental, between-subjects design	Questionnaires
Crutzen et al. (2013)	The Netherlands	To assess whether social presence may increase website use.	Healthy adults	Website	N/A	N = 458; Mean age (SD) = 49.0 (16.0); % Female = 50	Experimental, between-subjects design	Questionnaires

Cugelman et al. (2011)	UK	To explore the effect of persuasive and psychological design features to inform the development of online campaigns that seek to encourage health behaviour change.	Healthy adults	Internet-based interventions	N/A	N = 6028; % Female = 48	Systematic review, meta-analysis	N/A
Cussler et al. (2008)	US	To compare weight regain in women randomised to receive an online intervention and those randomised to self-directed weight maintenance.	Overweight and obese women	Website	12 months	N = 161; Mean age (SD) = 48.0 (4.4); % Female = 100,	RCT	Website logs
Danaher et al. (2006)	US	To describe initial patterns of participant exposure to ChewFree.com.	Smokeless tobacco users	Website	6 weeks	N = 2523	RCT	Website logs
Davies et al. (2012)	Australia	To assess the relationship between website engagement and intervention outcomes in 10,000 Steps Australia.	Healthy adults	Website	24 months	N = 348; % Female = 64	Cohort	Website logs
Dennison et al. (2014)	UK	To assess whether POWeR intervention usage was enhanced by the addition of brief telephone coaching.	Overweight individuals	Website	12 weeks	N = 786; Mean age (SD) = 44.0 (12.7); % Female = 80	RCT	Website logs
Donkin et al. (2011)	Australia	To describe methods used to assess adherence to e-therapy and to evaluate the association of adherence and intervention outcomes.	Physical illness and mental health	Technology-driven interventions	N/A	N = 34,465	Systematic review, narrative synthesis	N/A
Donovan et al. (2015)	US	To assess the efficacy of Wellness & Success among community and college students.	Alcohol and other drug users	Computer-based	120 minutes	N = 415; Mean age (SD) = 21.4 (2.2); % Female = 73	RCT	Website logs

Fang et al. (2013)	US	To develop an instrument to measure flow elements in computer gameplay.	N/A	Digital game	N/A	N/A	Instrument development and validation	Survey
Ferguson (2015)	UK	To examine adherence to online interventions for individuals with hearing loss.	Hearing loss	Web-based	4 weeks	N = 44; Mean age (SD) = 65.3 (5.7); % Female = 34	RCT	Website logs
Funk et al. (2010)	US	To examine website use patterns associated with long-term weight maintenance.	Overweight individuals at risk of cardiovascular disease	Website	30 months	N = 348; % Female = 63	RCT	Website logs
Geraghty et al. (2013)	US	To model attrition in a dual-language internet smoking cessation intervention.	English or Spanish speaking smokers	Internet intervention	4 weeks	N = 16430; Mean age (SD) = 36.2 (10.7); % Female = 47	RCT	Survey
Glasgow et al. (2011)	US	To characterise usage patterns in the My Path self-management website.	Individuals with Type 2 diabetes	Website	4 months	N = 270; Mean age = 60; % Female = 48	RCT	Website logs
Graham et al. (2013)	US	To determine whether smokers recruited during the New Year period differed on website utilisation rates compared with smokers recruited during other time periods.	Smokers	Website	3 months	N = 136; Mean age (SD) = 43.2 (12.3); % Female = 71	Secondary analysis of RCT data	Website logs

Habibovic et al. (2014)	The Netherlands	To assess characteristics of ‘completers’ and ‘non-completers’ in the WEBCARE intervention.	Patients with cardioverter defibrillators	Website	12 weeks	N = 146; Mean age (SD) = 58.2 (9.9); % Female = 18	RCT	Website logs
Haines-Saah et al. (2015)	Canada	To determine the feasibility of engaging young adults in a user-driven, online support forum for smoking cessation.	Smokers	Web-based	12 weeks	N = 60; Mean age = 21 (range: 19-24); % Female = 43	Cohort	Manual entry of website activities
Han et al. (2012)	US	To assess social and psychological characteristics predictive of different levels of engagement with an online support group.	Women with a diagnosis of breast cancer	Web-based	4 months	N = 231; Mean age = 51; % Female = 100	Pre- posttest	Website logs
Harmat et al. (2015)	Sweden	To assess the co-variation of subjective ratings of flow with cardiovascular and respiratory responses whilst playing a computer game.	Healthy adults	Digital game	N/A	N = 77; Mean age (SD) = 27.8 (5.4); % Female = 52	Experimental, within-subjects design	ECG recording, respiratory belt
Herbert et al. (2010)	Canada	To examine whether the Theory of Planned Behaviour and the Transtheoretical Model are able to explain adherence and attrition in an online intervention.	Chronic insomnia	Website	5 weeks	N = 94; % Female = 62	RCT	Questionnaires
Helander et al. (2014)	Finland	To assess factors associated with sustained use of The Eatery, a mobile app that promotes healthy eating.	Individuals interested in healthy eating	Smartphone app	N/A	N = 189,770	Cohort	Usage data

Henshaw et al. (2015)	UK	To explore motivations for uptake, engagement, and adherence to a computer-based auditory training programme.	Hearing loss	Computer-based	4 weeks	N = 44; Age range = 50-74	Randomised, quasi-crossover	Questionnaires, focus group
Hilvert-Bruce et al. (2012)	Australia	To examine whether non-completers drop out due to lack of efficacy and whether changes in delivery or clinician contact improve adherence.	Anxiety, depression	Online intervention	N/A	Study 1: N = 2107; Mean age (SD) = 40.1 (13.7); % Female = 64 Study 2: N = 1108; Mean age (SD) = 39.1 (13.6); % Female = 62 Study 3: N = 1090; Mean age (SD) = 40.1 (13.8); % Female = 64	Pre- posttest	Website logs
Hong et al. (2012)	Taiwan	To assess whether computer self-efficacy and 'competitive anxiety' are associated with flow.	Healthy adults	N/A	N/A	N = 101; % Female = 56	Cross-sectional	Survey
Horsch et al. (2015)	The Netherlands	To gain insight into strategies that enhance adherence to technology-mediated treatment.	Insomnia	Internet-based interventions	N/A	N = 2,961	Systematic review, meta-analysis; Qualitative	Face-to-face interviews, focus groups
Hsu & Lu (2004)	Taiwan	To identify predictors of users' acceptance of online games.	Healthy adults	Digital game	N/A	N = 233; % Female = 20	Cross-sectional	Survey
Hwang et al. (2011)	Taiwan	To explore the perceived usability of video games in an elderly population.	Elderly individuals (> 60 years)	Digital game	N/A	N = 60; % Female = 53	Qualitative	Interviews and observation

Irvine et al. (2015)	US	To evaluate the efficacy of FitBack.	Individuals with non-specific lower back pain	Mobile web app	8 weeks	N = 597; % Female = 58	RCT	Questionnaires
Jahangiry et al. (2014)	Iran	To assess adherence and attrition in a lifestyle intervention.	Individuals with metabolic syndrome	Website	6 months	N = 160; Mean age (SD) = 44.5 (10); % Female = 34	RCT	Attendance at follow-up assessment
Jennett et al. (2008)	UK	To assess whether immersion can be defined quantitatively.	Healthy adults	Digital game	N/A	N = 40; Mean age (SD) = 21.0 (3.5); % Female = 75	Experimental, between-subjects design	Questionnaires, eye tracking
Jennings (2000)	US	To describe theory and research from different disciplines relevant to creating engaging websites.	N/A	Website	N/A	N/A	Review, narrative synthesis	N/A
Johansson et al. (2015)	Sweden	To explore participants' experiences of non-adherence to Internet-delivered psychological treatment.	Generalised anxiety disorder	Website	8 weeks	N = 7; Mean age (SD) = 39.3 (17.1); % Female = 86	Qualitative	Face-to-face interviews
Kelders et al. (2012)	The Netherlands	To investigate whether particular intervention characteristics and persuasive design elements influence adherence to web-based interventions.	Health interventions	Web-based	N/A	N/A	Systematic review, narrative synthesis	N/A

Khadjesari et al. (2011)	UK	To determine the impact of incentives on follow-up rates in an online trial.	Alcohol users	Website	12 months	N = 7,935; Mean age = 38; % Female = 57	RCT	Questionnaires
Kim et al. (2013)	US	To test whether a novel mobile user engagement model may explain intention to engage.	Healthy adults	Smartphone	N/A	N = 297; % Female = 50	Cross-sectional	Survey
Klein et al. (2014)	The Netherlands	To assess the functioning of an intelligent mobile support system for therapy adherence and behaviour change.	Individuals with Type 2 diabetes, HIV, and/or cardiovascular disease	Smartphone app	N/A	N = 17	Pre- posttest	Survey
Kok et al. (2014)	The Netherlands	To examine user characteristics associated with adherence to the Mobile CT programme.	Depression	Website and mobile phone	8 weeks	N = 129	RCT, intervention arm	Website logs
Kuijpers et al. (2013)	The Netherlands	To explore the possible relevance of web-based interventions aimed at increasing empowerment and physical activity in individuals with chronic illness for cancer survivors.	Chronic illness	Web-based	N/A	N/A	Systematic review, narrative synthesis	N/A
Lefebvre et al. (2010)	US	To develop an instrument to measure engagement with health information.	Healthy adults	Website	N/A	N = 230; % Female = 60	Instrument development and validation	Questionnaires
Leslie et al. (2005)	Australia	To describe engagement and retention with a physical activity website in a workplace setting.	Healthy adults	Website	8 weeks	N = 655; Mean age = 43; % Female = 50	RCT, intervention arm	Website logs

Lieberman (2006)	US	To develop and evaluate the effect of a personified guide on adherence to an online alcohol reduction programme.	Alcohol users	Website	N/A	N = 288; Mean age (SD) = 36.0 (12.1); % Female = 31	RCT	Website logs
Lin & Wu (2014)	China	To assess the impact of SMS reminders on adherence to follow-up in digital health interventions.	Health interventions	Mobile phone	N/A	N = 12,783	Systematic review, meta-analysis	N/A
Liu et al. (2009)	Taiwan	To examine user acceptance of three kinds of streaming media (text, audio, and video) during online learning.	Healthy adults	Multimedia	N/A	N = 88	Experimental, between-subjects design	Survey
Ludden et al. (2015)	The Netherlands	To assess the impact of different design features on adherence to web-based wellbeing interventions.	Healthy adults	Web-based	N/A	N/A	Review, narrative synthesis	N/A
Mahmassani et al. (2010)	US	To examine user behaviour in a multiplayer online role-playing game.	Healthy adults	Digital game	N/A	N/A	Cohort	Game logs
Manwaring et al. (2008)	US	To assess whether adherence predicts outcomes in an online programme for the prevention of eating disorders.	Individuals with high levels of weight concern	Internet intervention	8 weeks	N = 209; % Female = 100	RCT, intervention arm	Website logs

Martey et al. (2014)	US	To examine the relationships among different measures of engagement.	Healthy adults	Digital game	N/A	Study 1: N = 280; Mean age = 21; % Female = 59 Study 2: N = 480; Mean age = 19.5; % Female = 65	Experimental, between-subjects design	Questionnaires, electro-dermal activity, mouse clicks, and mouse movement
McCabe & Price (2009)	Australia	To evaluate the dropout rate for an internet-based intervention for erectile dysfunction.	Erectile dysfunction	Website	12 weeks	N = 44; % Female = 0	RCT	Questionnaires
McCambridge et al. (2011)	UK	To determine whether differences in length and relevance of follow-up questionnaires have an impact on loss to follow-up.	Alcohol users	Website	12 months	N = 8,060	RCT	Questionnaires
McClure et al. (2013)	US	To explore the effect of four design features on engagement with an internet-based smoking cessation programme.	Smokers	Website	8 weeks	N = 1865; Mean age (SD) = 44.2 (14.7); % Female = 63	Multiphase optimization strategy trial	Website logs
Meischke et al. (2011)	US	To determine the characteristics of parents who engage with an internet-based health intervention for their children.	Parents to children with asthma	Website	6 months	N = 283	RCT, intervention arm	Website logs, survey

Miller, Cafazzo, & Seto (2014)	Canada	To examine effective use of gamification design principles in developing mHealth apps.	Chronic illness	Smartphone app	N/A	N/A	Review, narrative synthesis	N/A
Mohr et al. (2013)	US	To evaluate the efficacy of telephone coaching in improving adherence to MoodManager.	Depression	Website	12 weeks	N = 101	RCT	Website logs
Morris et al. (2015)	US	To introduce and evaluate a web-based, peer-to-peer cognitive reappraisal platform.	Depression	Website	3 weeks	N = 166; Mean age (SD) = 23.7 (5.3); % Female = 72	RCT	Website logs, questionnaires
Morrison & Doherty (2014)	UK	To conduct an exploration of the use of visualisations of log data to improve understanding of engagement with web-based interventions.	Depression	Website	N/A	N = 326	Secondary analysis of cohort data	Website logs
Morrison et al. (2014)	UK	To examine the effect of two different design features (tailoring and self-assessment) on engagement.	Mild bowel problems	Website	N/A	Study 1: N = 24; Median age = 25; % Female = 67	Qualitative	Interviews
						Study 2: N = 178; Mean age (SD) = 30.2 (11.7); % Female = 78	Partial factorial design	Website logs
Murray et al. (2013)	UK	To assess whether adherence and retention are related.	Alcohol users	Website	12 weeks	N = 7,932; Mean age (SD) = 38.0 (11.0); % Female = 57	Secondary analysis of RCT data	Website logs, questionnaires

Neve et al. (2010)	Australia	To describe the prevalence and predictors of dropout and non-usage attrition in a web-based weight loss programme.	Overweight individuals	Website	12 months	N = 9,599; Mean age (SD) = 35.7 (9.5); % Female = 86	Cohort	Website logs
Nicholas et al. (2010)	Australia	To explore reported reasons for non-adherence to an online psycho-education programme.	Bipolar disorder	Website	8 weeks	N = 39; % Female = 56	RCT; Qualitative	Website logs; Interviews
O'Brien & Toms (2008)	Canada	To conceptually and operationally define engagement with technology.	Healthy adults	Technology	N/A	N = 17; % Female = 59	Review, narrative synthesis; Qualitative	Face-to-face interviews
O'Brien & Toms (2010)	Canada	To develop an engagement scale.	Healthy adults	Website	N/A	Study 1: N = 440; % Female = 69 Study 2: N = 802; % Female = 70	Instrument development and validation	Questionnaires
Oh & Sundar (2015)	US	To explore the effect of two different interactivity types (modality and message) on website engagement.	Healthy adults	Website	N/A	N = 167; Mean age = 19.6; % Female = 58	Experimental, between-subjects design	Questionnaires
Oinas-Kukkonen & Harjumaa (2009)	Finland	To describe a framework for the design and evaluation of Persuasive Systems.	N/A	N/A	N/A	N/A	Review, narrative synthesis	N/A

Park et al. (2010)	US	To examine the effect of exposure to a pre-game story on the feeling of presence during gameplay.	Healthy adults	Digital game	N/A	Study 1: N = 30; % Female = 80 Study 2: N = 24; % Female = 58	Experimental, between-subjects design	Questionnaires
Parks (2014)	US	To outline important design considerations in online positive psychological interventions.	Healthy adults	Website	N/A	N/A	Review, narrative synthesis	N/A
Peels et al. (2012)	The Netherlands	To assess user characteristics associated with participation and attrition in web-based and print-based tailored physical activity interventions.	Aging population	Website	12 weeks	N = 1,729; Mean age = 48.3 % Female = 52	Cluster RCT	Questionnaires
Poirier & Cobb (2012)	US	To examine the association between social ties and engagement with a health and wellness online intervention.	Healthy adults	Website	4 weeks	N = 84,828; % Female = 84.	Cohort	Website logs
Postel et al. (2011)	The Netherlands	To examine attrition prevalence and pre-treatment predictors of attrition in a sample of open-access users of a Web-based program.	Problem drinkers	Website	3 months	Study 1: N = 780; Mean age (SD) = 47.5 (10.8); % Female = 54 Study 2: N = 144; Mean age (SD) = 45.8 (9.7); % Female = 58	Cohort (Study 1), RCT (Study 2)	Website logs

Richardson et al. (2010)	US	To measure the effect of adding online community features to an Internet-based walking program on attrition and average daily step counts.	Sedentary adults	Website	16 weeks	N = 324; Mean age (SD) = 52.0 (11.4); % Female = 66	RCT	Step counts, website logs
Richardson et al. (2013)	US	To examine the effectiveness of a Web-based smoking cessation intervention, to identify the most effective features, and to gain insight into who is most likely to use those features.	Smokers	Website	1 month	N = 1,033, % Female = 52	Cohort	Website logs
Ritterband et al. (2009)	US	To propose a model to help guide future development of online interventions and to predict and explain behavior change afforded by online interventions.	N/A	Internet interventions	N/A	N/A	Review, narrative synthesis	N/A
Sainsbury et al. (2015)	Australia	To assess the acceptability of an online intervention to improve diet adherence in coeliac disease and to examine the relationships with participant characteristics, attrition, and effectiveness.	Coeliac disease	Website	N/A	N = 189	RCT	Completion of follow-up assessment
Schønau-Fog & Bjørner (2012)	Denmark	To propose a method that can be used to empirically investigate the experience of wanting to continue playing.	Healthy adults	Digital game	N/A	N = 30	Qualitative	Interviews
Schubart et al. (2011)	US	To review what factors influence user engagement in Internet-based behavioral interventions for chronic illness.	Chronic illness	Internet-based interventions	N/A	N/A	Systematic review, narrative synthesis	N/A
Schwarzer & Satow (2012)	Germany	To predict smoking abstinence in internet users who engage with a virtual community.	Smokers	Virtual community	10 weeks	N = 13,174	Cohort	Website logs

Sharek & Wiebe (2014)	US	To investigate a novel technique for measuring video game engagement by capturing behavioral data without interfering with the main task.	Healthy adults	Digital game	N/A	N = 156; Mean age (SD) = 30.8 (10.2); % Female = 58	Experimental, between-subjects design	Game-clock clicks
Shaw et al. (2014)	US	To describe how fidelity recommendations may be applied in mobile phone interventions for weight loss.	Overweight individuals	Mobile phone	N/A	N = 261	Review, narrative synthesis	N/A
Short et al. (2015)	Australia	To propose a new model of user engagement that can be used to guide the development and evaluation of online behaviour change interventions.	N/A	Online interventions	N/A	N/A	Review, narrative synthesis	N/A
Stark et al. (2011)	US	To describe dietary self-monitoring rates among participants randomised to the intervention arms of two pilot studies.	Dialysis patients	Electronic diary	16 weeks	Study 1: N = 22; Mean age = 56; % Female = 40 Study 2: N = 26; Mean age = 52; % Female = 44	RCT, intervention arm	Website logs

Steinberg et al. (2014)	US	To examine patterns and predictors of self-monitoring adherence and the association between adherence and weight change in an online intervention.	African-American women with low income	Interactive voice response	12 months	N = 185; Mean age (SD) = 35.4 (5.5) % Female = 100	RCT	IVR completion
Strecher et al. (2008)	US	To determine whether engagement in a web-based smoking cessation intervention predicts 6-month abstinence, whether particular groups are more likely to engage, and whether particular components influence engagement.	Smokers	Website	6 months	N = 1,866; Mean age = 46.3; % Female = 60	Fractional factorial design with 16 arms	Website logs
Ubhi et al. (2015)	UK	To conduct a preliminary evaluation of the effectiveness of a novel smoking cessation smartphone application.	Smokers	Smartphone application	28 days	N = 1,170 % 16-29 years = 50.4 % 30-49 years = 45.4 % 50+ years = 4.2 % Female = 64.5	Observational prospective cohort	Automated recording logins, time spent, page views
VanDeMark et al. (2010)	US	To describe the characteristics of participants in the E-TREAT intervention, and to examine the characteristics that predict active engagement.	Substance use disorder	Website	3 months	N = 157; Mean age (SD) = 36.6 (9.7); % Female = 52	Cohort	Contact log

Van den Berg et al. (2006)	The Netherlands	To assess engagement with an Internet-based physical activity intervention with individual supervision.	Rheumatoid arthritis	Web-based	12 months	N = 82; Median age (IQR) = 49.5 (12.9); % Female = 76	RCT, intervention arm	Website logs
Vandelanotte et al. (2007)	Australia	To review outcomes of web-based physical activity interventions and to identify relationships of intervention components with behavioural outcomes.	Healthy adults	Website	N/A	N = 4,845	Systematic review, narrative synthesis	N/A
Voils et al. (2014)	US	To present approaches to inform intervention duration, frequency, and amount when 1) the researcher has no a priori expectation, and 2) when the researcher does have an a priori expectation.	N/A	N/A	N/A	N/A	Review, narrative synthesis	N/A
Wang et al. (2012)	US	To examine the mediating role of adherence to self-monitoring of diet and physical activity on weight loss in an online trial.	Overweight individuals	Web-based	12 months	N = 210; Mean age (SD) = 46.8 (9.0); % Female = 85	RCT	Usage data
Wanner et al. (2010)	Switzerland	To assess and compare user characteristics and adherence to Active Online in an open access context over time and between trial participants and open access users.	Healthy adults	Website	6 weeks	Study 1: N = 836; Mean age = 43.1; % Female = 75 Study 2: N = 5,084; Mean age = 38.4; % Female = 50	RCT (Study 1), cohort (Study 2)	Website logs

Webber et al. (2008)	US	To examine the relationships between motivation, adherence, and weight loss in an online behavioral weight-loss intervention.	Overweight individuals	Website	16 weeks	N = 66; Mean age (SD) = 50.1 (9.9); % Female = 100	RCT	Website logs
West & Michie (2016)	UK	To provide guidance on the development and evaluation of digital behaviour change interventions in healthcare.	N/A	N/A	N/A	N/A	Review	N/A
Weston et al. (2015)	UK	To investigate measurements of engagement using a health-based quiz app.	Healthy adults	Smartphone app	N/A	N = 29; Age range= 21-56; % Female = 59	RCT	Usage data
Whiteside et al. (2014)	US	To get user feedback on messaging content intended to engage suicidal individuals.	Suicidal individuals	Web-based	N/A	N = 34; % Female = 68	Cross-sectional	Survey
Zhou (2013)	China	To identify factors associated with the initial adoption of mobile games.	Healthy adults	Mobile phone	N/A	N = 231 % Female = 37	Cross-sectional	Survey

Appendix 3 – Online screening surveys (Study 2)

Online screening survey (all interested participants)

Question	Response Options
How old are you?	*Enter free text*
Do you smoke cigarettes daily?	(1) Yes (2) No
How often do you have a drink containing alcohol?	(0) Never (1) Monthly or less (2) 2 to 4 times a month (3) 2 to 3 times a week (4) 4 or more times a week
How many standard drinks containing alcohol do you have on a typical day when you are drinking?	(0) 1 to 2 (1) 3 to 4 (2) 5 to 6 (3) 7 to 9 (4) 10+
How often do you have six or more standard drinks on one occasion?	(0) Never (1) Less than monthly (2) Monthly (3) Weekly (4) Daily or almost daily
Do you live in or near London?	(1) Yes (2) No
Do you own an iPhone or an Android smartphone with Internet access capable of running apps?	(1) Yes (2) No
Would you consider using a smartphone app to help you stop/cut down on your smoking?	(1) Yes (2) No
Would you consider using a smartphone app to help you cut down on your drinking?	(1) Yes (2) No

Online survey to assess demographic characteristics (all eligible participants)

Question	Response Options
What is your gender?	(1) Male (2) Female
Do (did) you work as an employee or are (were) you self-employed?	(1) Employee (2) Self- employed with employees (3) Self-employed/freelance without employees
How many people work (worked) for your employer at the place where you work (worked)? OR How many people do (did) you employ?	(1) 1 to 24 (2) 25 or more
Do (did) you supervise any other employees? (A supervisor or foreman is responsible for overseeing the work of other employees on a day-to-day basis)	(1) Yes (2) No
Please tick one box to show which best describes the sort of work you do. If you are not working now, please tick a box to show what you did in your last job.	(1) Modern professional occupations (teacher, nurse, physiotherapist, social worker, welfare officer, artist, musician, police office, software designer) (2) Clerical and intermediate occupations (secretary, personal assistant, clerical worker, office clerk, call centre agent, nursing auxiliary, nursery nurse) (3) Senior managers or administrators (finance manager, chief executive) (4) Technical and craft occupations (motor mechanic, fitter, inspector, plumber, printer, tool maker, electrician, gardener, train driver) (5) Semi-routine manual and service occupations (postal worker, machine operative, security guard, caretaker, farm worker, catering assistant, receptionist, sales assistant) (6) Routine manual and service occupations (HGV driver, van driver, cleaner, porter, packer, sewing machinist, messenger, labourer, waiter/waitress, bar staff) (7) Middle or junior managers (office manager, retail manager, bank manager, restaurant manager, warehouse manager, publican)

	(8) Traditional professional occupations (solicitor, medical practitioner, scientist, civil/mechanical engineer)
What is your ethnic group?	(1) English/Welsh/Scottish/Northern Irish/British (2) Irish (3) Gypsy or Irish Traveller (4) Any other White background (5) White and Black Caribbean (6) White and Black African (7) White and Asian (8) Any other Mixed/Multiple ethnic background (9) Indian (10) Pakistani (11) Bangladeshi (12) Chinese (13) Any other Asian background (14) African (15) Caribbean (16) Any other Black/African/Caribbean background (17) Arab (18) Any other ethnic group
Have you made an attempt to stop smoking in the past 12 months?	(1) Yes (2) No
Which of the following best describes you?	(1) I don't want to stop smoking (2) I think I should stop smoking but I don't really want to (3) I want to stop smoking but I haven't thought about when (4) I REALLY want to stop smoking but I don't know when I will (5) I want to stop smoking and hope to soon (6) I REALLY want to stop smoking and intend to in the next 3 months (7) I REALLY want to stop smoking and intend to in the next month
Have you ever used a smartphone app to help you quit smoking?	(1) Yes (2) No
How many cigarettes do you smoke per day?	*Enter free text*
How soon after waking do you usually smoke your first cigarette?	(0) 61+ minutes (1) 31-60 minutes (2) 6-30 minutes (3) <5 minutes
How often did you experience urges to smoke in the past 24 hours?	(0) Not at all (1) A little of the time (2) Some of the time (3) A lot of the time (4) Almost always (5) All the time

Have you made an attempt to cut down on drinking alcohol in the past 12 months?	(1) Yes (2) No
Which of the following best describes you?	(1) I don't want to cut down on drinking alcohol (2) I think I should cut down on drinking alcohol but don't really want to (3) I want to cut down on drinking alcohol but haven't thought about when (4) I REALLY want to cut down on drinking alcohol but I don't know when I will (5) I want to cut down on drinking alcohol and hope to soon (6) I REALLY want to cut down on drinking alcohol and intend to in the next 3 months (7) I REALLY want to cut down on drinking alcohol and intend to in the next month
Have you ever used a smartphone app to help you cut down on your drinking?	(1) Yes (2) No
How often did you experience urges to drink in the past 24 hours?	(0) Not at all (1) A little of the time (2) Some of the time (3) A lot of the time (4) Almost always (5) All the time
When was the last time you downloaded an app, if ever?	(1) Today or yesterday (2) In the last week (3) In the last month (4) In the last 3 months (5) In the last 6 months (6) More than 6 months ago
How frequently do you use the apps on your smartphone, if at all?	(1) Daily (2) Weekly (3) Monthly (4) Never
Do your friends and family ask for your advice or help in using smartphone apps?	(1) Yes (2) No
I use my smartphone to:	(1) Check my e-mail (2) Find out what my friends are doing on Facebook (3) Get information via Twitter (4) Navigate using Google Maps or similar tools (5) Read the news (6) Research things to purchase (7) Download and play games (8) Download and use health/fitness apps

Appendix 4 – Topic guides and verbal instructions (Study 2)

Pre-session interview

1. Can you tell me about an app that you are using regularly? Why do you think that you are using it regularly?

2. Have you ever used a health or fitness app? Can you tell me about it?

3. What do you think a smoking cessation/alcohol reduction app should provide or do?

4. After the first half of the interviews, another question was added:

5. What is your identity as a smoker/drinker?

Think aloud

Verbal instructions

"During this session, you will be given two smartphone-based tasks to complete. I would like to emphasise that this is not a test; I am interested in the tasks themselves, not your performance. I would like you to complete the tasks whilst 'thinking aloud'. This means that I would like you to complete the tasks, and while you do so, try to say everything that goes through your mind. I would like you to pretend that you are at home and try to forget that I am here.

Thinking aloud usually feels a bit strange at first, as it is an unusual task. Don't worry about it, most people find it a bit unnatural at first, but quickly get used to it! We will start off with a practice task to make sure that you feel comfortable. I would like you to change the ring tone on your smartphone whilst trying to say everything that goes through your mind."

Tasks

1. I would like you to imagine that you are at home. Please find an app that you think will be engaging enough to help you quit or cut down on your smoking/drinking. Please use the App Store/Google Play to search for apps whilst thinking aloud.

2. Imagine that you have selected a smoking cessation/alcohol reduction app that you would like to try. Please download one of the free apps that you think will be engaging enough to help you quit or cut down on your

smoking/drinking. Please complete the baseline questions and explore the app whilst thinking aloud.

Post-session interview

1. I noticed that you mentioned that you thought that [...] was ... Can you tell me a bit more about that?
2. I noticed that you made a comment about [...]. Can you elaborate on that?
3. How do you understand the term ‘engagement’ in the context of apps?
4. Do you think that the app that you chose to download was engaging? Why/why not?
5. Do you think that you would find the app/those particular features engaging longer term? Why/why not?
6. Do you think that you might use the app that you have downloaded after leaving this session? Why/why not?

After the first half of the interviews, the following questions were added:

7. You mentioned that you thought that [...] was ... How do you think that feature would fit into your daily life?
8. How important is it for you to be able to relate to the app’s content?
9. How do you think engaging with the app would help you stop/cut down on your smoking/drinking?
10. How do you think that [...] would help you stop/cut down on your smoking/drinking?

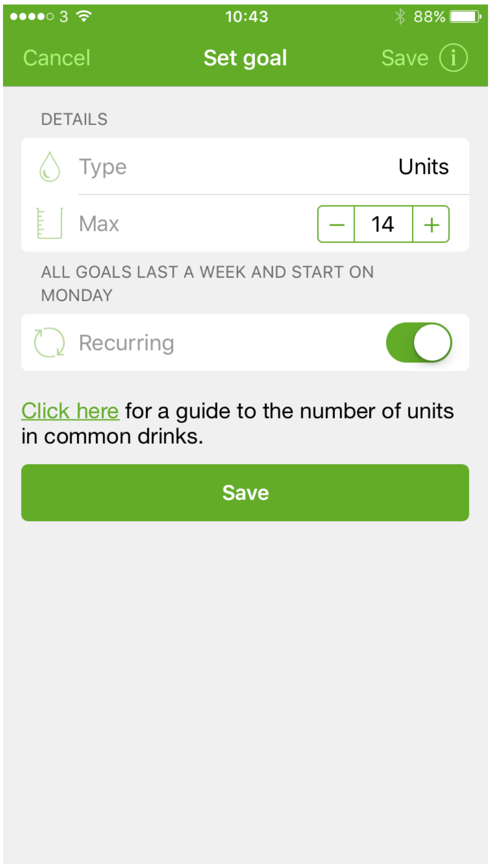
Appendix 5 – Additional quotations (Study 2)

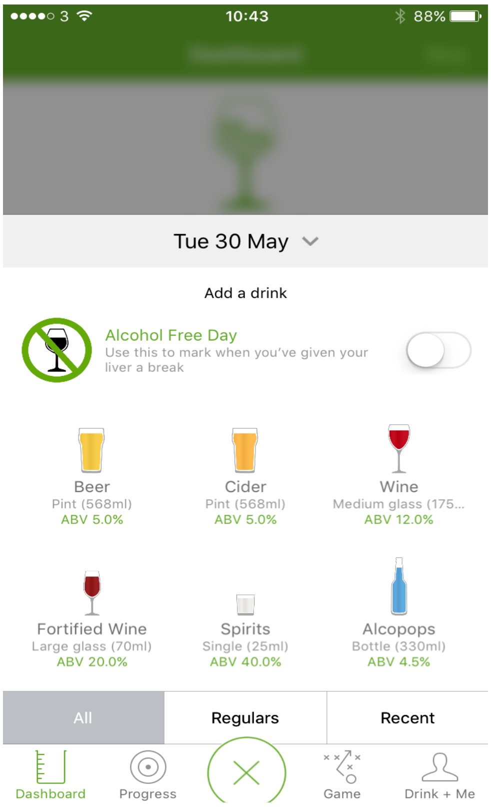
Research Question	Theme	Example Excerpts
1. What factors shape smokers' and drinkers' choice of apps?	*The immediate look and feel of the app*	"...they were sort of dark and black and had some sort of neon lighting on it, and it just didn't look very inviting, whereas that first one was actually really colourful and bright, and quite modern, and had illustrations…" – D3
		"...it looks kind of inviting, with all the colours and exclamations and stuff." – D5
		"I liked the logo because it was green, it looked minimalistic, so I thought that maybe the app will be easy to use, and not chaotic, just easy to use." – S5
		"The pictures look very scientific, not very user-friendly at all." – D1
		"It looks a lot more simple to use, and a simple user face." – S4
	Social proof	"It hasn't had that many downloads, only 100, which seems quite low, not many people have used this." – D4
		"None of these apps have any ratings, so it's really hard to know what people are thinking of them, because I genuinely just go on the ratings." – S2
		"So here's the thing, this is what I normally do, I would just look at the stars, because other people have done this." – S6
		"It's also by […], so it just seems more trustworthy. I don't know the other ones." – D7
		"...the fact that it was by the […] made me want to go and look at it." – S2
	Realistic and relevant titles	"There's one called […] but it's got the word 'alcoholic' in it so I'm thinking it's probably not for me." – D2
		"“Dependence – quit tobacco and alcohol”. That sounds a bit hip, directed towards younger people than me." – D6
		“Obviously, with that, maybe, I'd known what I was getting, which probably would have been more suitable, but, “Sober Time,” it didn't seem like it was completely off topic, so it didn't mislead me in that way.” – D10
		"OK, this one is one of the condescending ones. “We are your motivation” kind of thing." – S3
		“I think they could probably have chosen a better name. I didn't like the whole “now” thing." – S6

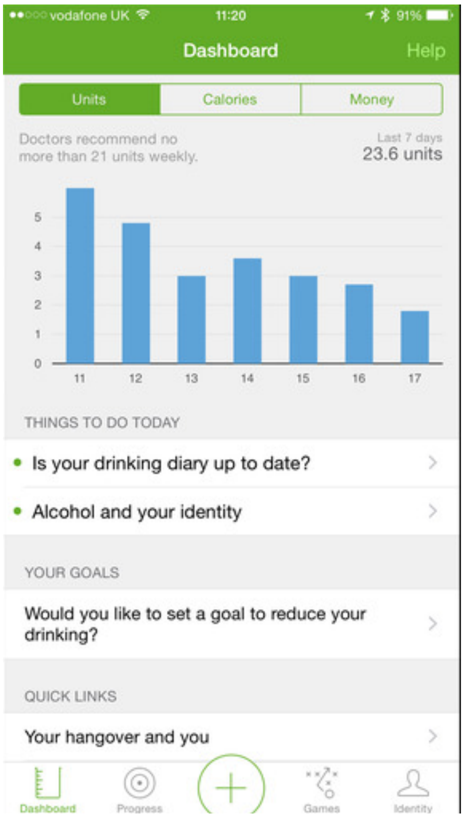
2. What factors are judged to be important for engagement?	*Features that enhance motivation*	"I liked that. That was a good idea. You have a goal to aim for, and then it's saying: "You've reached it" or "You haven't reached it."" – D4 "...then you have these goals as well, and then obviously if you do have a cigarette then, you kind of ruin it…" – S5 "...it just tallies up how much you're smoking, which is good, because just keeping track of something will automatically help you reduce…" – S2 "Rewards to keep me motivated to give all this information in the journal. Just to keep using it." – S6 "I mean, I can see why these achievement things are useful, but for me, they don't really bother me at all. I guess it's trying to motivate you to get all these achievements, but for me, I just kind of don't give a s*** about that." – D1
	Features that enhance autonomy	"...if there's an option saying: "How often do you want to be notified", I think would be quite handy as well." – D4 "I want to be able to handle the app myself, and feel that it's support for me, not taking over." – D6 "I like having control, but then, I probably will forget to use things unless it gives me a notification." – S9 "Because that 14 units of alcohol is just an irrelevance. It doesn't apply to anyone at all." – D9 "I do want to quit cold turkey, and then I'd probably want one of those apps, but in the meantime, I would like an app that would also be for cutting down smoking, because that's obviously helping too." – S5
	Features that enhance personal relevance	"...the app doesn't personalise enough. Maybe in the registration bit, they could perhaps ask you what you're interested in mostly." – D7 ""Heavy alcohol consumption is linked to a number of cancers, such as..." OK, thank you. Now you're just scaring me. I don't know if I like them telling me that." – D6 "...not only how you're destroying everything with the smoking, because obviously that doesn't go in, but maybe positive things will…" – S3 "It comes back to what I was saying about the guidelines coming across as preachey or whatever." – D1 "...it has very non-judgmental language in it, which is really good." – S10

	Features that enhance credibility	"Yeah, I'm not bothered about data concerns, or whatever, I'm quite happy to just give my information away." – D1
		"I just don't think I like to give my details to any random app that I've searched for, so, like I said, if someone recommended something to me, but yeah, I probably don't like to give my information to anything more than I have to, and definitely not a random app with bad reviews..." – D3
		"If you say something inaccurate, stupid, or lazy, you're going to lose your credibility." – S6
		"I don't know whether they've just thought that could help or if they've actually conducted some research where people have said that this has helped, tips in the moment." – S9
	Consistency with online and offline social preferences	"If I'm going to do this, it would be about what I was doing, I don't really mind if there's a strong community of people doing exactly the same thing." – D10
		"I don't know, I don't think I would want other people to know that I'm trying to reduce my drinking, personally." – D4
		"I don't see very many tips on that, the social aspect and how to get over the social aspect." – S10
		"...that's what I see as the psychology of quitting. If you are quitting by telling everyone that you're quitting, it puts so much pressure on you." – S3
		"“Post progress on Facebook or Twitter.” I wouldn't be interested in that, I'm more of a private person." – S6

Appendix 6 – Screen shots of the *Drink Less* app (Study 3)






vodafone UK
11:20
91%
Dashboard
Help
Units
Calories
Money
Doctors recommend no
more than 21 units weekly.
Last 7 days
23.6 units
5
4
3
2
1
0
11
12
13
14
15
16
17
THINGS TO DO TODAY
Is your drinking diary up to date?
Alcohol and your identity
YOUR GOALS
Would you like to set a goal to reduce your
drinking?
QUICK LINKS
Your hangover and you
Dashboard
Progress
Games
Identity

Appendix 7 – Online screening survey (Study 4)

Question	Response Options
Please enter your Prolific ID.	*Enter free text*
What is your gender?	(1) Male (2) Female
How old are you (in years)?	*Enter free text*
Are you currently residing in the United Kingdom?	(1) Yes (2) No
What kind of job do you have?	1) Manual 2) Non-manual 3) Other (e.g. student, unemployed, retired)
Do you own an iPhone capable of running iOS v.8.0 or higher (i.e. iPhone 4S or later models)?	(1) Yes (2) No
Are you willing to download and explore an alcohol reduction app?	(1) Yes (2) No
How often do you have a drink containing alcohol?	(1) Never (2) Monthly or less (3) 2 to 4 times a month (4) 2 to 3 times a week (5) 4 or more times a week
How many standard drinks containing alcohol do you have on a typical day when you are drinking?	(1) 1 or 2 (2) 3 or 4 (3) 5 or 6 (4) 7, 8 or 9 (5) 10 or more
How often do you have six or more standard drinks on one occasion?	(1) Never (2) Less than monthly (3) Monthly (4) Weekly (5) Daily or almost daily
How often during the last year have you found that you were not able to stop drinking once you had started?	(1) Never (2) Less than monthly (3) Monthly (4) Weekly (5) Daily or almost daily
How often during the last year have you failed to do what was normally expected from you because of drinking?	(1) Never (2) Less than monthly (3) Monthly (4) Weekly (5) Daily or almost daily
How often during the last year have you needed a first drink in	(1) Never (2) Less than monthly (3) Monthly

the morning to get yourself going after a heavy drinking session?	(4) Weekly (5) Daily or almost daily
How often during the last year have you had a feeling of guilt or remorse after drinking?	(1) Never (2) Less than monthly (3) Monthly (4) Weekly (5) Daily or almost daily
How often during the last year have you been unable to remember what happened the night before because you had been drinking?	(1) Never (2) Less than monthly (3) Monthly (4) Weekly (5) Daily or almost daily
Have you or someone else been injured as a result of your drinking?	(1) No (2) Yes, but not in the last year (3) Yes, during the last year
Has a relative or friend or doctor or another health worker been concerned about your drinking or suggested you cut down?	(1) No (2) Yes, but not in the last year (3) Yes, during the last year
Which of the following best describes you?	(1) I don't want to cut down on drinking alcohol (2) I think I should cut down on drinking alcohol but don't really want to (3) I want to cut down on drinking alcohol but haven't thought about when (4) I REALLY want to cut down on drinking alcohol but I don't know when I will (5) I want to cut down on drinking alcohol and hope to soon (6) I REALLY want to cut down on drinking alcohol and intend to in the next 3 months (7) I REALLY want to cut down on drinking alcohol and intend to in the next month

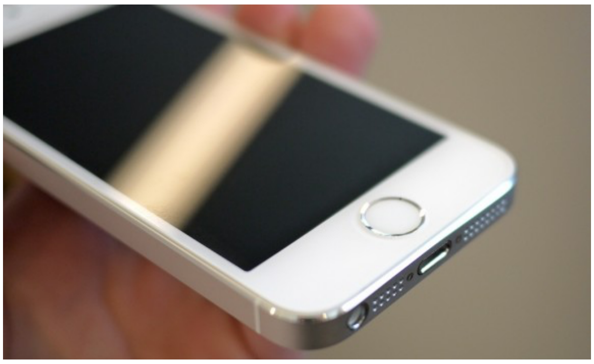
Appendix 8 – Task instructions for Prolific participants (Study 4)

1. Please go to the Apple App Store and download the *Drink Less* app:

https://itunes.apple.com/gb/app/drink-less-help-drinking-less/id1020579244?mt=8

2. Complete the onboarding process. Please spend some time exploring the app, in the same manner as you would explore any other new app. **Please ensure that you click 'Allow' when you are asked whether you would like to receive push notifications, as this is how you will receive a link to the final survey.**

3. When you have finished exploring the app, please press your iPhone's home button. This will trigger a push notification asking you to fill out a brief survey. Please click on the notification and complete the survey. Please make sure that you enter your Prolific ID!



Appendix 9 – Recruitment materials (Study 5)

Do you drink 2 pints of beer/glasses of wine 2-3 times a week? Are you interested in using a smartphone app to reduce your drinking?

We are looking for participants to take part in a focus group study to help us learn more about how to design engaging smartphone apps for alcohol reduction. Joining the study will involve a visit to University College London's main campus, where you will be asked to participate in a focus group discussion with other participants.

The visit will take approximately 2 hours and you will receive a £20 gift voucher as compensation for your time.

You are eligible to take part if you:

- **Are 18 years or older**
- **Own an iPhone or Android smartphone**
- **Live in or near London**
- **Consume alcohol regularly**
- **Are interested in using a smartphone app to reduce your drinking**
- **Have previously used a health or fitness app**

Please be advised that these are not the only eligibility criteria and that you must complete an online screening survey before we can invite you to take part in the study. The screening survey can be accessed here: tinyurl.com/y82qkd6k

If you are interested in taking part, please contact olga.perski.14@ucl.ac.uk for more information.

This study is conducted by: Olga Perski, PhD candidate, Department of Clinical, Educational and Health Psychology, and is supervised by Professor Susan Michie, Professor Ann Blandford and Professor Robert West. The study has received ethical approval from UCL's Research Ethics Committee (Project ID: UCLIC/1213/015).

Appendix 10 - Topic guide for focus groups (Study 5)

1. Can you tell us a bit more about why you ranked [insert feature here] highly?

2. Can you tell us a bit more about why you ranked [insert feature here] as less important for engagement?

3. How do you think that [insert feature here] would help you to engage with the app?

4. Could you give a concrete example of [insert feature here] that would help you to engage with the app?

Appendix 11 – Additional quotations (Study 5)

Themes	Quotations
1. Lack of trust and guidance as initial barriers	"… you'd only use it as soon as you set up the app and that's it. When you know how to work your way around the app and what features there are, and that's it. Once you know how it works, then you don't need it." – P1, focus group
	"…it shows you around the app first, shows you what it's capable of doing, the little things it's got going on, and then it's like: "Hey! Let's start this"." – P5, focus group
	"I would like to know that the app is from a credible source before I even contemplate using it, so this would be a must to begin." – P24, online sample
2. Motivational support	"I feel that if you decide to carry out a task, you need to be in control of it, because ultimately, that's your goal that you're setting, and you want to have a sense of ownership or control of whatever you want to achieve. You feel more responsible for how you carry out your goals." – P2, focus group
	"I think if rewards were present I would be more likely to use the app on a regular basis." – P33, online sample
	"I am competitive - really enjoy challenges within my network." – P45, online sample
3. Benefit and usefulness	"The information you enter, you want to get something back from it, just to increase your improvement, or whatever you want to achieve." – P2, focus group
	"I think to feel you're getting something else out of it, it will make you think not having alcohol can still be a good thing even if that's what you want…" – P26, online sample
	"I think that's the only way you're going to get it to work. It's got to be a two-way thing between you and the app to achieve your goals and to make sure that information that you put in is correct as well. Otherwise the whole thing is going to be a waste of time." – P1, focus group

4. Adaptability	"In the beginning, you might want to set a goal, but once the app gets more information about you, it might suggest a further goal, so it's more challenging…" – P4, focus group
	"I think a pop-up notification could be really useful around like 6 or 5pm, which is when people get off work or off uni…" – P9, focus group
	"Any app I engage with needs to meet my individual needs. Apps that do not have flexibility will not be used by me, regardless of their function." – P76, online sample
5. Sparking users' interest	"You want it to be interesting to you. You want it to have surprising features." – P4, focus group
	"I feel like if there was games and quizzes to do it would encourage me to use it more as I like doing these." – P110, online task
	"It just keeps you want to engage with it, or it doesn't make you want to keep away, it makes you want to go back to it because it's actually quite good, and more features are unlocking as you're going on. You don't realise this, and this is happening…" – P5, focus group
6. Relatedness	"It just gets a lot of people in the same situation as you. You don't feel so alone." – P2, focus group
	"Would be able to get help and support from similar people going through the same problems." – P41, online sample
i. Perceived social stigma	"I don't like the idea of putting something personal and having lots of people seeing it." – P6, focus group
	"See, I'm quite private about things like drinking more than I want to, which is sort of where I'm up to, or I'm a lot healthier now, but there are few people that I would share that with." – P7, focus group
ii. Fear of social comparison	"…people like me that would be quite shy about my friends knowing my success or failure, day by day. I think that might be quite overwhelming…" – P7, focus group
	"To me, it's the self-betterance. I'm trying to improve myself, so I'd rather do it just me rather than someone saying: "I used to do this..." – P5, focus group

Appendix 12 – Recruitment materials (Study 6)

Do you drink alcohol on a regular basis? Are you interested in using a smartphone app to reduce your drinking?

We are looking for participants to take part in a study to help us learn more about how people use alcohol reduction apps in their daily lives.

Taking part in this study will involve:

- A briefing interview, conducted in person at University College London (60-90 minutes)
- Daily use of an alcohol reduction app for a period of 28 days
- The completion of two brief surveys (~1 minute) per day for 28 days, sent to you via text message
- A debriefing interview, conducted over the phone (30 minutes)

Providing that you complete at least 70% of the brief surveys and participate in the debriefing interview, you will receive a £60 gift voucher as compensation for your time and effort.

You are eligible to take part if you:

- Are a fluent English speaker
- Are 18 years or older
- Own an iPhone capable of running iOS 8.0 software or higher (i.e. iPhone 4S or later models)
- Live in or near London and are willing to come into University College London for the briefing interview
- Regularly consume alcohol (approximately 2 pints of beer/glasses of wine, 2-3 times per week)
- Are interested in using an app to reduce your drinking
- Are willing to set a goal to reduce your drinking
- Are willing to use an alcohol reduction app daily for 28 days, recognising that there may be occasional days where you will not use it

If you are interested in taking part, please contact olga.perski.14@ucl.ac.uk for more information and to complete the screening survey.

This study is conducted by: Olga Perski, PhD candidate, Department of Clinical, Educational and Health Psychology, and is supervised by Professor Susan Michie, Professor Ann Blandford and Professor Robert West. The study has received ethical approval from UCL's Research Ethics Committee (Project ID: UCLIC/1617/004/Staff Blandford HFDH).


Call For Participants
Understanding the real-world use of an alcohol reduction app
28 day(s) to complete
£60 shopping voucher
Interview; Daily surveys for 28 days
Gower St, Bloomsbury, London WC1E 6BT, UK
University College London
We are looking for participants to take part in a study to help us learn more about how people use an alcohol reduction app in their daily lives. This study involves: 1) a briefing interview, conducted in person at University College London (60-90 minutes); 2) daily use of an alcohol reduction app for a period of 28 days; 3) the completion of two brief surveys (~1 minute) per day for 28 days, sent to you via text message; 4) a debriefing interview, conducted over the phone (30 minutes).
Find out more online
Poster printed on 28/06/2018
Study expires on 28/07/2018
More info
by scanning the QR code or visiting the URL
www.cfp.cc /X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3
cfp.cc/X8RHR3

Appendix 13 – Online screening survey (Study 6)

Question	Response Options
What is your gender?	(1) Male (2) Female
How old are you (in years)?	*Enter free text*
What kind of job do you have?	1) Manual 2) Non-manual 3) Other (e.g. student, unemployed, retired)
Do you live in or near London?	(1) Yes (2) No
Do you own an iPhone capable of running iOS v.8.0 or higher (i.e. iPhone 4S or later models)?	(1) Yes (2) No
Are you willing to download and explore an alcohol reduction app?	(1) Yes (2) No
Have you ever used an alcohol reduction app? If so, which one?	(1) Yes (2) No *Enter free text*
Are you interested in using an app to reduce your drinking?	(1) Yes (2) No
Are you willing to set a goal to reduce your drinking?	(1) Yes (2) No
Are you willing to come into University College London for a briefing interview?	(1) Yes (2) No
Are you willing to use an alcohol reduction app daily for 28 days? (We recognise that there may be occasional days where you will not use it)	(1) Yes (2) No
Are you willing to respond to text messages two times per day for 28 days?	(1) Yes (2) No
Are you willing to participate in a debriefing interview at the end of the study, conducted over the phone?	(1) Yes (2) No
How often do you have a drink containing alcohol?	(1) Never (2) Monthly or less (3) 2 to 4 times a month (4) 2 to 3 times a week (5) 4 or more times a week
How many standard drinks containing alcohol do you have	(1) 1 or 2 (2) 3 or 4

on a typical day when you are drinking?	(3) 5 or 6 (4) 7, 8 or 9 (5) 10 or more
How often do you have six or more standard drinks on one occasion?	(1) Never (2) Less than monthly (3) Monthly (4) Weekly (5) Daily or almost daily
How often during the last year have you found that you were not able to stop drinking once you had started?	(1) Never (2) Less than monthly (3) Monthly (4) Weekly (5) Daily or almost daily
How often during the last year have you failed to do what was normally expected from you because of drinking?	(1) Never (2) Less than monthly (3) Monthly (4) Weekly (5) Daily or almost daily
How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session?	(1) Never (2) Less than monthly (3) Monthly (4) Weekly (5) Daily or almost daily
How often during the last year have you had a feeling of guilt or remorse after drinking?	(1) Never (2) Less than monthly (3) Monthly (4) Weekly (5) Daily or almost daily
How often during the last year have you been unable to remember what happened the night before because you had been drinking?	(1) Never (2) Less than monthly (3) Monthly (4) Weekly (5) Daily or almost daily
Have you or someone else been injured as a result of your drinking?	(1) No (2) Yes, but not in the last year (3) Yes, during the last year
Has a relative or friend or doctor or another health worker been concerned about your drinking or suggested you cut down?	(1) No (2) Yes, but not in the last year (3) Yes, during the last year
Which of the following best describes you?	(1) I don't want to cut down on drinking alcohol (2) I think I should cut down on drinking alcohol but don't really want to (3) I want to cut down on drinking alcohol but haven't thought about when

	(4) I REALLY want to cut down on drinking alcohol but I don't know when I will (5) I want to cut down on drinking alcohol and hope to soon (6) I REALLY want to cut down on drinking alcohol and intend to in the next 3 months (7) I REALLY want to cut down on drinking alcohol and intend to in the next month

Appendix 14 – Text messages (Study 6)

Twice-daily text message

Hi [insert name here]!

Please answer the following questions:

a) How motivated are you currently to reduce your drinking? (1 = not at all; 7 = extremely)

b) How useful do you currently think the *Drink Less* app is for you? (1 = not at all; 7 = extremely)

c) How many drinks containing alcohol have you had in the past 12 hours?

d) To what extent do you currently have time for the *Drink Less* app? (1 = I have lots of time for the app; 7 = I don't have any time for the app)

Please enter your responses as follows: a=X; b=X; c=X; d=X

When response is received

Thank you for your responses!

Message if response is not in expected format

Hi [insert name here]! It appears that your responses are not in the expected format. Please enter your responses as follows: a=X; b=X; c=X; d=X

Thank you!

Weekly message about EMA completion rate

Hi [insert name here]! Thank you for completing the first week of the study. You have responded to X out of 14 text messages. Keep up the good work!

Hi [insert name here]! Thank you for completing the second week of the study. You have responded to X out of 28 text messages. Keep up the good work!

Hi [insert name here]! Thank you for completing the third week of the study. You have responded to X out of 42 text messages. Keep up the good work!

Hi [insert name here]! Thank you for completing the first week of the study. You have responded to X out of 56 text messages. Keep up the good work!

Appendix 15 – Published peer-reviewed articles

TBM

SYSTEMATIC REVIEWS

CrossMark

Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis

Olga Perski,[1] Ann Blandford,[2] Robert West,[3] Susan Michie[1]

[1]Department of Clinical, Educational and Health Psychology, University College London, 1-19 Torrington Place, London, WC1E 6BT, UK

[2]UCL Interaction Centre, University College London, 66-72 Gower Street, London, WC1E 6EA, UK

[3]Cancer Research UK, Health Behaviour Research Centre, Department of Epidemiology and Public Health, University College London, 1-19 Torrington Place, London, WC1E 6BT, UK

Correspondence to: O Perski olga.perski.14@ucl.ac.uk

Cite this as: *TBM* 2017;7:254–267
doi: 10.1007/s13142-016-0453-1

Abstract

"Engagement" with digital behaviour change interventions (DBCIs) is considered important for their effectiveness. Evaluating engagement is therefore a priority; however, a shared understanding of how to usefully conceptualise engagement is lacking. This review aimed to synthesise literature on engagement to identify key conceptualisations and to develop an integrative conceptual framework involving potential direct and indirect influences on engagement and relationships between engagement and intervention effectiveness. Four electronic databases (Ovid MEDLINE, PsycINFO, ISI Web of Knowledge, ScienceDirect) were searched in November 2015. We identified 117 articles that met the inclusion criteria: studies employing experimental or non-experimental designs with adult participants explicitly or implicitly referring to engagement with DBCIs, digital games or technology. Data were synthesised using principles from critical interpretive synthesis. Engagement with DBCIs is conceptualised in terms of both experiential and behavioural aspects. A conceptual framework is proposed in which engagement with a DBCI is influenced by the DBCI itself (content and delivery), the context (the setting in which the DBCI is used and the population using it) and the behaviour that the DBCI is targeting. The context and "mechanisms of action" may moderate the influence of the DBCI on engagement. Engagement, in turn, moderates the influence of the DBCI on those mechanisms of action. In the research literature, engagement with DBCIs has been conceptualised in terms of both experience and behaviour and sits within a complex system involving the DBCI, the context of use, the mechanisms of action of the DBCI and the target behaviour.

Keywords

Engagement, Digital, Behaviour change interventions, eHealth, mHealth, Conceptual framework, Systematic review

Implications

Practice: The use of a shared conceptual framework for engagement with digital behaviour change interventions (DBCIs) should promote more rapid advances in developing methods to improve it.

Policy: A shared conceptualisation of engagement with DBCIs can be used to help policymakers and commissioners to set standards against which to evaluate DBCIs.

Research: The proposed conceptual framework can be used to generate testable hypotheses about how to improve engagement.

Electronic supplementary material
The online version of this article (doi:10.1007/s13142-016-0453-1) contains supplementary material, which is available to authorized users.

INTRODUCTION

A substantial number of Internet-connected adults use some forms of digital technology to monitor or modify their health: estimates vary between 20 and 80% [1–3]. Digital behaviour change interventions (DBCIs), defined as "…a product or service that uses computer technology to promote behaviour change" [4], can, for example, be delivered through computer programs, websites, mobile phones, smartphone applications (apps) or wearable devices. Evidence suggests that DBCIs can help people change a range of different health behaviours, including smoking [5, 6], alcohol consumption [7], weight management [8], physical activity [9] and self-management of chronic conditions [10]. Some form of "engagement" with DBCIs is assumed to be important for their effectiveness [11]. A positive association between engagement and, for example, smoking cessation, weight loss and increased fruit and vegetable intake has been observed [12–14]. To date, we have not achieved a shared understanding of how to usefully conceptualise and operationalise engagement with DBCIs. This systematic review, which follows the *Cochrane Collaboration's Handbook of Systematic Reviews of Interventions* [15], examines how engagement has been construed and measured in the behavioural science, computer science and human-computer interaction (HCI) literatures and uses this to propose an integrative definition and conceptual framework of engagement with DBCIs that can be used to generate predictions and explanations of empirical observations.

The design of DBCIs requires knowledge of intervention content, delivery, interface design and computer programming, which have traditionally been informed by separate scientific disciplines, such as behavioural science, computer science and HCI. Scientific disciplines are characterised by accumulating a body of specialist knowledge and developing a specific terminology concerned with the particular object of research [16]. Due to the multifaceted structure of DBCIs, an interdisciplinary approach, where knowledge from multiple disciplines is harnessed to develop a shared viewpoint, is required to develop a useful conceptualisation of engagement in this context [17].

Engagement has traditionally been conceptualised differently across the behavioural science, computer science and HCI literatures, which might be due to the different epistemologies subscribed to, the differing research contexts and the different objectives pursued. In the computer science and HCI literatures, engagement has traditionally been conceptualised as the subjective experience of flow, a mental state characterised by focused attention and enjoyment [18]. This kind of conceptualisation might have emerged as a result of the focus on entertainment and usability of interactive technology. In the behavioural science literature, engagement has typically been conceptualised as "usage" of DBCIs, focusing on the temporal patterns (e.g. frequency, duration) and depth (e.g. use of specific intervention content) of usage [19, 20]. This kind of conceptualisation has emerged due to the observation that while many download and try DBCIs, sustained usage is typically low [21–24]. Henceforth, two working definitions of engagement as used in the computer science and HCI literatures ("engagement as flow") and the behavioural science literature ("engagement as usage") are used to scope the space within which this review is conducted.

Although existing systematic reviews have assessed whether particular DBCI features (e.g. tailoring, reminders) are associated with higher engagement [25, 26] and whether engagement is associated with intervention effectiveness [11], it is not possible to synthesise results from these reviews or to draw any conclusions regarding the shape of the function (e.g. linear, non-linear) relating engagement with intervention outcomes due to the use of incomparable definitions of engagement [11]. In order to reduce fragmentation of research efforts, it would be useful to develop a shared understanding of how to conceptualise and operationalise engagement with DBCIs.

A conceptual framework can been defined as "a system of concepts, assumptions, and expectations, and the presumed relationships among them" [27]. Previous conceptual frameworks of engagement have proposed multiple interacting factors (e.g. social support, sensory appeal, ease of use) that influence "engagement as flow" or "engagement as usage" [28–30]; however, these frameworks are either not derived from empirical observations or draw only on literature from one of many interrelated scientific disciplines. For example, the framework proposed by O'Brien and Toms [28], notwithstanding its grounding in empirical observations, drew only on research from the technology literature and focused on "engagement as flow" without any links to behaviour change. Conversely, the framework by Ritterband and colleagues [29] focused on "engagement as usage" and was derived from behavioural science theory only. The model proposed by Short and colleagues [30] attempted to integrate both theoretical predictions and empirical findings from the behavioural science, persuasive design and technology literatures but did not do so in a systematic manner. Although the ontology of behaviour change interventions proposed by West and Michie provides a starting point for organising and representing DBCIs, engagement constitutes one of many important components and is hence not examined in detail [4]. It is therefore not possible to determine whether existing frameworks of engagement sufficiently explain real-world events, or whether important aspects are missing.

The aims of this review are threefold; the second and third build on output from the first:

1. To synthesise past work on engagement, addressing the following research questions:

(a) How has engagement been defined in the selected literatures?
(b) How has engagement been measured?
(c) What factors have been found or hypothesised to influence engagement?
(d) What are the proposed relationships between engagement and intervention effectiveness?

2. To develop an integrative definition of engagement with DBCIs and specify how it can be measured.
3. To develop a conceptual framework of the direct and indirect influences on engagement with DBCIs and the proposed relationships between engagement and intervention effectiveness.

METHODS

The *Cochrane Handbook of Systematic Reviews of Interventions* [15] and the *Guidance for Undertaking Reviews in Health Care* [31] were used to inform the development of the search strategy, identify inclusion criteria, select studies and extract the data. Principles from critical interpretive synthesis (CIS) were used to inform the data synthesis [32]. As CIS is one of the few methods available that affords the synthesis of qualitative and quantitative data, it was deemed to be the most suitable method. CIS is useful when a review seeks to identify a definition of a phenomenon, as it aims to produce a higher-order structure or conceptual framework ("synthesising argument"), which is grounded in the concepts ("synthetic constructs") identified in the reviewed articles [32]. CIS does not propose a formal method for critically appraising the quality and methodological rigour of included studies but recognises that the critical evaluation and integration of disparate forms of evidence is essentially a product of the "authorial voice" [33]. The evidence is critiqued on the basis of the implicit assumptions underlying the methodological decisions made in the reviewed articles. Hence, the quality of the evidence is considered in the development of the synthetic constructs, with the consideration based on the authors' judgements. Principles of CIS have previously been employed in reviews of the health literature [34–36].

Criteria for considering studies for this review

All types of study designs were included except position papers. All types of information sources were included except articles that were not peer-reviewed or not available in English. Studies with adult participants (i.e. aged 18 years or older) were included, as it was expected that different factors might influence engagement in children and adult populations due to different cognitive abilities [37]. Studies specifically targeting participants with cognitive impairment or intellectual disabilities were excluded for the same reason. DBCIs and digital interventions targeting individuals with mental health or chronic physical health conditions were included as no a priori reason suggesting that engagement should be conceptualised differently across the included topic areas could be identified. Interventions were excluded if they did not incorporate any digital component as part of the intervention itself (i.e. face-to-face delivery only) or if the technology was used solely as a tool to deliver measurement surveys. Studies involving recreational or educational digital games or multimedia software (e.g. software involving animations, sound and text) were included providing that engagement was discussed or measured. For the conceptualisation of "engagement as flow", the games or technology did not need to be related to behaviour change. The primary outcome was definitions of engagement with DBCIs, digital games or multimedia software expressed either implicitly or explicitly. Secondary outcomes included proposed direct and indirect influences on engagement, measures of engagement and associations between engagement and intervention effectiveness expressed either implicitly or explicitly.

Search methods for the identification of studies

Electronic searches

A structured search of the following electronic databases was conducted in November 2015: Ovid MEDLINE (1946–November 2015), PsycINFO (1806–November 2015), ISI Web of Knowledge (1900–November 2015) and ScienceDirect (1900–November 2015). Search terms were piloted and refined to achieve a balance between sensitivity, i.e. retrieving a high proportion of relevant articles, and specificity, i.e. retrieving a low proportion of irrelevant articles [15]. An academic librarian was consulted for the validation of the databases and the final search terms. Terms were searched for in titles and abstracts as free text terms or as index terms (e.g. Medical Subject Headings) where appropriate (see Electronic Supplementary Material 1).

Searching for other resources

Articles from adjacent fields not immediately or obviously relevant to the research questions were identified through expertise within the review team [32]. The Association for Computing Machinery Digital Library (a repository for conference proceedings) and relevant journals (i.e. *Journal of Medical Internet Research*, *Journal of the American Medical Informatics Association*, *Telemedicine and e-Health*) were hand searched, and reference chaining was employed to identify additional articles of interest [15, 32].

Data collection and analysis

Selection of studies

Articles identified through the electronic and hand searches were merged using EndNote X7 [38] to ensure consistency. Duplicate records were removed. Two researchers independently screened (i) titles, (ii) abstracts and (iii) full texts of the identified articles against the predefined eligibility criteria [15]. Any disagreements were resolved through discussion and by consulting a third researcher if necessary. Inter-rater reliability was assessed based on two coding categories (i.e. inclusion versus exclusion) after the full text screening phase with the prevalence- and bias-adjusted kappa (PABAK) statistic, which controls for chance agreement [39]. The following cutoffs were used: 0.40–0.59 indicates fair agreement, 0.60–0.74 indicates good agreement and >0.75 indicates high agreement [15].

Data extraction and management

A pro-forma was developed by the first author to extract information about the study setting, participant characteristics, study design, data collection method and study findings [32]. The pro-forma was piloted on a sample of included articles to ensure that relevant information was captured [15]. A second researcher independently checked the pro-forma for accuracy and completeness [31]. Due to limited resources, a single reviewer completed the data extraction.

Quality appraisal

CIS suggests the prioritisation of seemingly relevant articles rather than favouring particular study methodologies [40]. Judgements about the relevance and underlying assumptions of articles were made by the first author and were incorporated into the data synthesis [32].

Data synthesis

Based on the principles from CIS, the data synthesis comprised the following steps:

1. Concepts identified in the full texts of included articles were labelled with codes by the first author. The research questions were used as a top-down coding frame; fragments of text explicitly or implicitly referring to definitions of engagement, measures of engagement, influences on engagement or associations between engagement and intervention effectiveness were coded.
2. A subsample of codes was selected through random sequence generation (https://www.random.org/) for validation by an independent researcher to increase rigour [41]. Disagreements were discussed until consensus was reached.
3. Synthetic constructs (i.e. concepts that explain similar themes) were developed from the codes, and relationships between synthetic constructs were specified by the first author.
4. The synthetic constructs and the proposed relationships between constructs were validated by an independent researcher. Disagreements were discussed until consensus was reached.
5. Two synthesising arguments (i.e. an integrative definition and its measurement, and a conceptual framework) were developed based on the synthetic constructs by the first author.
6. The synthesising arguments were refined through discussion between all co-authors.

Results

Summary of search results

The electronic database search yielded 925 published articles. After removing duplicates, 560 articles remained for screening. A PABAK score of 0.88 was achieved after the full text screening phase, indicating high inter-rater reliability [15]. Due to this reliability score, the additional 31 information sources were screened by a single reviewer. Of the 140 full texts screened, 117 met the inclusion criteria and were included in the data synthesis. Six qualitative studies, 27 reviews, 2 mixed methods studies and 82 quantitative studies were included (see Fig. 1). Characteristics of the included studies are described in Electronic Supplementary Material 2.

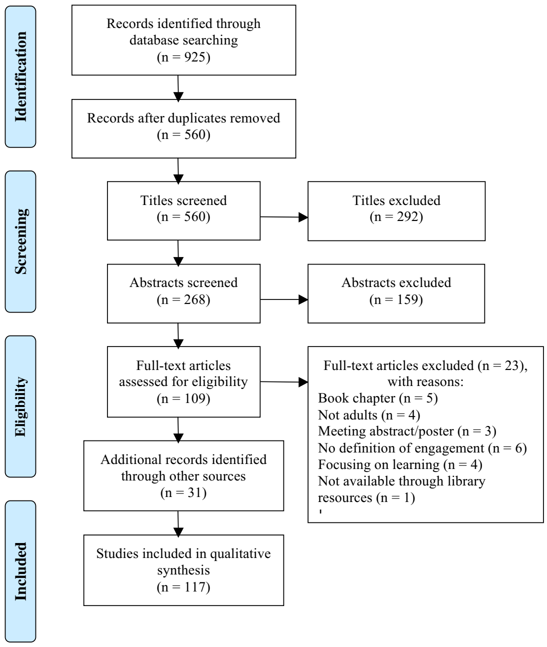


Fig 1 | PRISMA flow diagram of the study selection process [42]

How has engagement been defined in the literature?

The following two synthetic constructs were developed: "engagement as subjective experience" and "engagement as behaviour".

Engagement as subjective experience

Engagement has been conceptualised as the *subjective experience* that emerges in the momentary interaction with a system [18, 28, 43]. This kind of conceptualisation was only identified in the computer science and HCI literatures. Similarities can be found between engagement and the state of "flow", described as a mental state characterised by focused attention, intrinsic interest and enjoyment, balance between challenge and skill, and temporal dissociation (i.e. losing track of the passage of time) [18, 44–47]. Similarities can also be found between engagement and the state of "immersion" within digital gaming, characterised by cognitive absorption, the willingness to direct emotions towards an activity and feeling cutoff from reality [43, 48–51]. As conceptual overlap was observed between these experiential qualities, the authors propose that they can be grouped under the following cognitive and emotional states: attention, interest and affect.

Engagement as behaviour

The majority of articles reviewed from the behavioural science literature conceptualised engagement in *behavioural* terms, suggesting that it is identical to the usage of a DBCI or its components. Engagement has further been described as the extent of usage over time [19, 52], sometimes referred to as the "dose" obtained by participants or "adherence" to an intervention [25, 53, 54], determined by assessing the following subdimensions: "amount" or "breadth" (i.e. the total length of each intervention contact), "duration" (i.e. the period of time over which participants are exposed to an intervention), "frequency" (i.e. how often contact is made with the intervention over a specified period of time) and "depth" (i.e. variety of content used) [20, 53]. In the computer science and HCI literatures, engagement has been conceptualised as the degree of involvement over a longer period of time [55], sometimes referred to as "stickiness" [56]. A distinction has also been made between "active" and "passive" engagement; while the former involves contributing to the intervention through posting in an online discussion forum, the latter involves reading what others have written without commenting, also known as "lurking" [57]. Engagement has also been conceptualised as a process of linked behaviours, suggesting that users move dynamically between stages of engagement, disengagement and re-engagement [28]. As conceptual overlap was observed between these definitions, the authors propose that engagement involves different levels of usage over time.

Development of an integrative definition of engagement

An integrative definition of engagement with DBCIs was developed through the merging of overlapping conceptualisations as outlined above, in addition to the integration of the two overarching synthetic constructs. The following two-part definition is therefore proposed:

"*Engagement with DBCIs is (1) the extent (e.g. amount, frequency, duration, depth) of usage and (2) a subjective experience characterised by attention, interest and affect*".

Engagement is conceptualised as a multidimensional construct: the behavioural dimensions of engagement are underpinned by the user's subjective experience of what it feels like to be engaged with a DBCI. Engagement is considered to be a dynamic process that is expected to vary both within and across individuals over time.

How has engagement been measured?

The following two synthetic constructs were developed: "subjective measures" and "objective measures".

Subjective measures

In research settings, self-report questionnaires have frequently been used to measure engagement with digital games and DBCIs [51, 58–67]. Qualitative approaches, such as interviews or think aloud methodology, have been employed to gain a better understanding of the nature of users' experiences of engagement with digital games and DBCIs [60, 68, 69].

Objective measures

Automatic tracking of use patterns, including number of logins, time spent online and the amount and type of content used during the intervention period, was the most commonly used measure of engagement in the behavioural science literature [11, 19, 20, 26, 44, 70–82]. Physiological measures including cardiac activity, respiratory depth [62] and electro-dermal activity [65], and psychophysical measures, such as eye tracking [51], have been used to measure engagement in the computer science and HCI literatures.

Measures relating to the integrated conceptualisation of engagement

Based on the literature synthesis, we suggest that all facets of engagement proposed in the integrative definition of engagement can in principle be measured or inferred through the following: (1) user-reported interaction with the DBCI through self-report questionnaires, interview studies or think aloud studies; (2) automated recording of DBCI use (e.g. logins, page views); and (3) recording of physiological or psychophysical correlates of DBCI interaction.

What factors have been hypothesised or found to influence engagement?

The following two synthetic constructs were developed: “context” and “DBCI”. Context was subdivided into “population” and “setting.” DBCI was subdivided into “content” and “delivery.” Relationships between constructs were specified.

Context

Population

Psychological characteristics–Motivation was found to be positively associated with engagement across many studies, with none indicating a negative association [20, 68, 83–87]. As the available evidence is correlational in nature, the direction of influence cannot be assumed. It has been hypothesised that the relationship between motivation and engagement might be U-shaped; those who are least and most motivated to, for example, quit smoking, are hypothesised to disengage quickly from DBCIs due to failed and successful behaviour change, respectively [19].

Expectations are thought to be influential in that users are hypothesised to engage more if there is a match between their expectations and the goal of the DBCI [49, 73, 86, 88, 89]. Prior experiences of using other websites or apps, or of having tried face-to-face counselling (which may or may not have worked), might shape users’ expectations of what DBCIs can provide [90].

Mental health, including low mood, anxiety and stress, has been found to be negatively associated with engagement [68, 73, 87, 91–96]. A negative association with mental health was mainly observed in studies of DBCIs targeting individuals diagnosed with a mental health condition but was also observed in physical activity [68] and weight loss [94] interventions. Similarly, *experience of well-being* or believing that one does not need to work on certain issues has been found to be negatively associated with engagement [92].

Need for cognition, defined as the tendency to process large amounts of information [11, 30, 57, 88, 97], and *self-efficacy* to execute a given behaviour [83, 98, 99] were found to be positively associated with engagement.

Personal relevance, which refers to the extent to which a DBCI is perceived to apply to the individual and their particular situation, has been hypothesised to positively influence engagement [69, 78, 100–104]. Results from interview studies indicate that participants believe that lack of personal relevance is a sufficient reason for dropping out from intervention trials [86, 92, 95, 105].

Demographic characteristics–Age [20, 57, 63, 68–70, 73, 76, 79, 91, 95, 96, 99, 106–111], *gender* [20, 69, 73, 90, 95, 100, 101, 110, 111], *education* [20, 69, 91, 92, 96, 99, 106, 107, 109, 110, 112], *employment* [91, 92, 107] and *ethnicity* [57, 106] were found to be significantly associated with engagement. There was a trend towards a positive association between engagement and older age, higher educational attainment and being a woman; however, as no meta-analysis was conducted, a conclusion about the size and direction of influence cannot be drawn. *Computer literacy*, or confidence using the Internet, has been found to be positively associated with engagement [11, 20, 98, 99, 106, 108, 113]. However, as none of the included studies adequately measured baseline computer skills in their designs, a firm conclusion cannot be drawn.

Physical characteristics–Physique, including baseline weight and the presence of comorbidities, was found to be negatively associated with engagement [68, 70, 71, 91–94, 106, 112].

Setting

The *social* and *physical* environments in which a DBCI is used, have been hypothesised to influence engagement [4, 29, 30]. The social environment includes culture (e.g. prevailing norms), commercial environment, media and social cues. The physical environment includes financial resources, material resources, time pressure, physical cues, location, the healthcare system and policy. *Time* [86, 92, 93, 114] and *access* to hardware or the Internet [30, 115] have been hypothesised to be positively associated with engagement.

DBCI

Content

DBCIs that include particular *behaviour change techniques* (BCTs), such as action plans [78], goal setting [116], feedback [59] and self-monitoring tools [78], have been found to be associated with higher engagement [78]. *Rewards* and *incentives* have been hypothesised [26, 100, 101, 117] or found [118] to positively influence engagement; however, evidence from trials in which the presence of rewards or incentives has been manipulated is scarce.

Social support features, referring to features that facilitate the receipt of social support, were found to positively influence engagement [76, 82, 119–124]. Features that decrease the feeling of loneliness or that increase motivation through competition with others include online discussion forums, gamification elements such as leaderboards that show users where they rank in a gamified system, and peer-to-peer contact [125, 126]. Evidence indicates that DBCIs that provide access to such features are successful in getting users who report lower social support at baseline to engage [57, 127]; however, participants who reported higher levels of social support at baseline were found to be more likely to engage with the social elements of DBCIs across a few studies [68, 86, 91, 96].

Reminders have been hypothesised [117, 128, 129] or found to positively influence engagement; results from a meta-analysis indicate a positive effect of reminders on engagement [130]. However, receiving too many reminders may have a negative effect on engagement due to "e-mail fatigue" [69].

Delivery

Mode of delivery, which includes face-to-face, telephone, text message, smartphone app, website and mass media delivery, has been hypothesised to influence engagement with DBCIs [4].

Professional support features, which include features that enable remote contact with a clinician via e-mail, telephone or text messages, have been found to positively influence engagement with DBCIs [20, 25, 26, 63, 68, 70, 73, 77, 88, 90, 95, 120, 131–134]. However, results from a randomised controlled trial (RCT) of a web-based weight loss intervention in which some participants received coaching calls from a nurse indicated that participants in the coaching arm were more likely to drop out around the time of the first coaching session, suggesting a negative influence of professional support features in particular situations [70].

Control features, referring to features that make users feel that they are in control of and are free to make choices about how to interact with a DBCI, have been hypothesised [51, 119] or found [52, 74, 110] to positively influence engagement. For example, results from an RCT in which participants either received content all at once or sequentially over a period of weeks suggest that participants were more likely to disengage when the content was delivered sequentially [110]. Tunnelled interventions (i.e. those that lead users through a number of predetermined steps) have been found to generate more page views compared with self-paced ones [74]. However, this may be an artefact of making users click through a pre-specified number of pages in order to progress through the DBCI.

Novelty, generated by regular content updates, has been found to positively influence engagement through preventing boredom [25, 26]. However, there might be a trade-off between novelty and programme *complexity*; it has been hypothesised that participants will disengage if the intervention is perceived as too long or overly complicated [26, 68, 73, 88, 131, 135, 136]. It has been hypothesised that the presence of too many features may compromise a DBCI's *ease of use* [19], referring to whether or not it feels natural for the user to operate an interactive system. Ease of use has been hypothesised to positively influence engagement [71, 100, 137].

The *personalisation* or tailoring of content has been hypothesised [26, 52, 68, 72, 80, 103, 106, 110, 113, 119, 120, 138] or found [19, 20, 66] to positively influence engagement. *Interactivity*, referring to a two-way flow of information between a DBCI and its user, has been hypothesised [28, 48, 50, 66, 78, 100, 139] or found [19] to positively influence engagement.

Message tone, which refers to the terminology and wording used to communicate health messages [92, 101], and *narrative* [43, 50, 65, 103, 125, 140], referring to the presence of a storyline, have been hypothesised to positively influence engagement. Furthermore, *challenge* [61, 100, 141], *aesthetics and design* [120, 139, 142, 143] and *credibility features* [68, 73], referring to features that inculcate a feeling of trust, *familiarity* [97, 139, 144], and the provision of *guidance* or tutorials [68, 106, 145] have been hypothesised to positively influence engagement with DBCIs.

What are the proposed relationships between engagement and the effectiveness of DBCIs?

The following four synthetic constructs were developed to explain the proposed relationships between engagement and the effectiveness of DBCIs: "mechanisms of action", "unmeasured third variable", "optimal dose" and "effective features".

Mechanisms of action

Mechanisms of action proposed to mediate the effect of engagement with DBCIs on intervention effectiveness [4] include increased knowledge, motivation, affect management, cognitive restructuring, skill building [29], comprehension and practice of programme content, and increased self-efficacy [19]. A further distinction has been made between "intervention receipt", which refers to the extent to which participants understand and can perform the skills taught, and "enactment of intervention skills", which refers to the extent to which participants use these skills [146, 147]. It has also been hypothesised that mechanisms of action, such as accountability to a healthcare practitioner and relatedness to other individuals, might positively influence engagement with DBCIs [68, 77, 86, 96].

Unmeasured third variable

An *unmeasured third variable*, such as higher baseline motivation or self-efficacy, may be responsible for the observed association between increased engagement and positive DBCI outcomes. Alternatively, those who engage with DBCIs might simply be more inclined to behave healthily in general [11]. It has also been argued that the *target behaviour* itself might influence engagement [148]. For example, smokers who relapse might be more likely to stop engaging with the DBCI, while those who successfully manage their cravings might be more likely to continue engaging with the DBCI.

Optimal dose

Optimal dose refers to a pre-defined level of engagement at which specific DBCIs are effective. It has been hypothesised that the receipt of an optimal dose may explain the relationship between engagement and intervention effectiveness but that the optimal dose for particular DBCIs may vary depending on user characteristics [70, 113].

Effective features
The use of specific intervention features has been found to be associated with better DBCI outcomes [70]. It has been suggested that there may be a mismatch between features that participants choose to engage with frequently and *effective features* that are causally linked to intervention outcomes [104]. For example, although users may enjoy engaging with a particular feature (e.g. filling out a food diary), thus using it frequently, use of a less popular feature (e.g. "getting support" tools) might be more strongly associated with intervention outcomes, such as weight loss [70].

Development of a conceptual framework of engagement with DBCIs
The final aim of the review was to develop a conceptual framework specifying potential direct and indirect influences on engagement and relationships between engagement and intervention effectiveness. As the framework proposed by Ritterband and colleagues [29] and the ontology proposed by West and Michie [4] explicitly linked engagement to behaviour change, we drew on these to structure our conceptual framework, mapping the other existing frameworks onto it. Additional factors identified in the reviewed literature not otherwise specified were also mapped onto the conceptual framework.

We propose a conceptual framework in which engagement with a DBCI influences the target behaviour through specific mechanisms of action; box 4, box 1, box 3 and box 2, respectively. Content has been found to directly influence engagement with DBCIs; box a. Delivery has been hypothesised to directly influence engagement with DBCIs; box b. The context and the target behaviour are hypothesised to directly influence engagement; box 5 and box 3, respectively. Mechanisms of action are hypothesised to indirectly influence engagement; box 2. The population (e.g. demographic, physical and psychological characteristics) has been found to directly influence engagement with DBCIs; box c. The setting has been hypothesised to directly influence engagement; box d. Engagement is hypothesised to be indirectly influenced by the moderating influence of the context on the influence of the DBCI; box 4, box 5 and box 1, respectively. Figure 2 shows this schematically. Hypothesised influences are marked with stars.

DISCUSSION
An integrative conceptualisation of engagement with DBCIs has been developed; engagement is defined here as a multidimensional construct which can be measured through self-report questionnaires, verbal reports, automatic recording of DBCI use or recording of psychophysical manifestations. A conceptual framework was developed, which suggests that the context of use influences engagement with DBCIs either directly or indirectly by moderating the influence of the DBCI on engagement. Mechanisms of action might indirectly influence engagement and the target

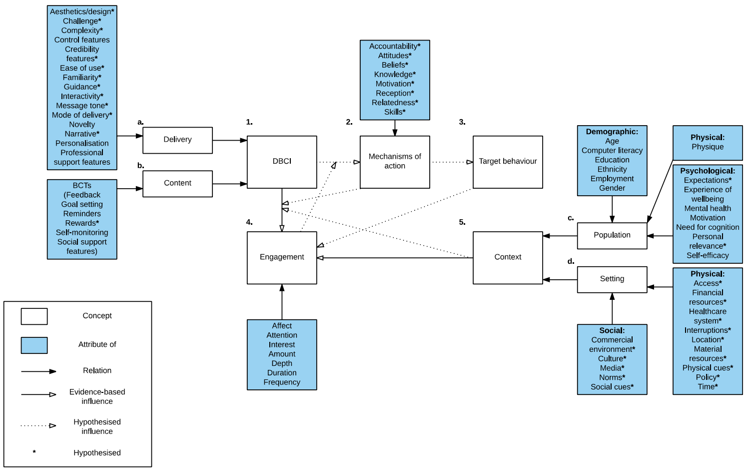


Fig 2 | Conceptual framework of direct and indirect influences on engagement with DBCIs. Transparent boxes indicate concepts. Concepts can be defined as abstract ideas that are derived from either direct or indirect evidence [149]. Blue boxes indicate attributes of concepts. Attributes can be defined as properties that characterise a concept [150]. Solid black arrows indicate relationships between concepts and attributes. Arrows with transparent heads indicate an influence of a concept.

behaviour might directly influence engagement with DBCIs, suggesting the presence of a positive feedback loop. The proposed relationships between engagement and intervention effectiveness are tentative, as these have not been studied extensively.

The suggested behavioural and experiential dimensions of engagement can in principle be measured or inferred in every instance of a DBCI. The content, structure, length and design of specific DBCIs tend to vary, and hence, the relevance of the different dimensions of engagement will vary accordingly. Although the intended frequency, amount, duration and depth of use might be set to “1” in a one-off intervention, the individual parameters are still present and measureable. Thus, the proposed definition of engagement allows for direct comparison across different kinds of DBCIs by including multiple dimensions of engagement at its core. This has been lacking in previous conceptualisations. Evidence of higher engagement coupled with evidence of, for example, enjoyment of using a DBCI is hypothesised to predict greater DBCI effectiveness. If this is the case, the proposed definition of engagement should provide a means of generalising findings from particular DBCIs to other similar DBCIs. It may not be possible to evaluate the usefulness of the proposed definition prior to empirical work [151].

Although some self-report questionnaires designed to measure engagement demonstrate good validity and reliability [64, 152], these typically rely on measuring engagement after, as opposed to during, the event. However, the advent of new technologies allows self-reports of engagement to be measured in real-time rather than through paper-and-pencil questionnaires [153]. Although physiological measures have been used to measure engagement, notably in the HCI literature, associations between physiological and self-reported measures of engagement are weak [65]. The nature of these associations should be investigated further.

Previous conceptual frameworks have been based on theoretical predictions only or have been derived from the literature within one scientific domain [4, 28–30]. In contrast, our conceptual framework is derived from theoretical predictions and empirical observations within multiple, interrelated disciplines. This endeavour was facilitated by the use of principles from CIS, which allowed the combination of a diverse set of research findings. The proposed conceptual framework of engagement is a synthesis of existing ontologies, frameworks and models and incorporates factors not previously included. The novel components in our framework are as follows: “mental health”, “experience of well-being”, “familiarity”, “guidance” and “narrative”. The negative association between poor mental health and engagement might be explained by the observation that those with poor mental health (e.g. depression) typically experience decreased self-efficacy to, for example, stop smoking or lose weight [154, 155]. Experience of well-being might be negatively associated with engagement due to being related to the belief that one does not need any support. Familiarity with the design of DBCIs and guidance might positively influence engagement because familiar examples, design conventions or stepped how-to-use guides may inculcate feelings of comfort and ease of use. A narrative might draw users in, increasing their interest and enjoyment. Moreover, this review identified a trend towards a positive association between engagement and older age, higher educational attainment and being a woman, which merits further investigation. Although these demographic characteristics have been included in existing frameworks of engagement, the direction of influence has not been previously discussed. Through the use of a systematic, interdisciplinary approach, the proposed conceptual framework offers a comprehensive overview of the factors that may influence engagement with DBCIs and hence provides a starting point for reducing the observed fragmentation of research findings.

LIMITATIONS

The lack of evidence supporting the claim that setting of use (e.g. culture, social norms, physical cues, healthcare pathway) directly influences engagement with DBCIs constitutes a limitation. This might either reflect the search terms used or indicate that this has not been investigated in the literature; we cannot distinguish between these explanations. There was also a lack of evidence in support of the claim that the context of use (i.e. setting and population) may moderate the influence of the DBCI on engagement. For example, the setting of use may vary depending on the mode of delivery (e.g. computer versus mobile phone). Hence, the DBCI might indirectly influence engagement through determining the setting of use; while computers may predominantly be used at home or in a clinic, mobile phones might mainly be used on the go, which may influence the amount or depth of engagement. Future research should test this hypothesis. Another limitation is that no formal quality assessment of the included articles was conducted. However, this was in line with the chosen method, which suggests that the articles should be judged on the basis of their relevance to the research question rather than their methodological rigour. This method was selected due to the conceptual nature of the research questions. A further limitation is that the data extraction and literature synthesis were conducted by a single reviewer, potentially introducing bias. Finally, the end date for the literature search (i.e. November 2015) constitutes a limitation; with the pace of technological advances and the proliferation of digital health research, it is likely that relevant literature has since been published.

IMPLICATIONS AND AVENUES FOR FUTURE RESEARCH

The proposed integrative definition and conceptual framework of engagement with DBCIs have implications for clinical practice: the use of a shared terminology and measurement techniques will ensure more rapid advance in understanding engagement with DBCIs and developing methods to improve it. A shared conceptualisation of engagement can be used to help policymakers and commissioners to set evaluation standards for DBCIs. Moreover, the proposed conceptual framework can be used to generate testable hypotheses about how to improve engagement with DBCIs. For example, according to the conceptual framework, the presence of rewards might influence engagement with a DBCI due to increased motivation. This hypothesised link between rewards, motivation and engagement can be tested using an experimental design. Future avenues for research include the assessment of what dimensions of engagement (e.g. attention, interest, affect, amount, duration, frequency, depth) are most strongly associated with intervention effectiveness, whether it is possible to establish benchmarks for the optimal dose of engagement across different kinds of DBCIs and whether the context of use influences engagement with DBCIs.

CONCLUSION

Engagement with DBCIs is conceptualised here in terms of both experience and behaviour. Engagement may be influenced by the DBCI itself, the context of use, mechanisms of action of the DBCI and the target behaviour.

Acknowledgements: The authors would like to thank Jacqui Smith, librarian at University College London, for helping to build the search strategy; Nicola Newhouse for validating the coding frame and commenting on an early draft; and Holly Walton for helping to screen articles. Olga Perski is a Ph.D. candidate funded by a grant from Bupa under its partnership with University College London. Susan Michie is part funded by grants from Cancer Research UK and NIHR's School for Public Health Research. Robert West is funded by Cancer Research UK. The funders played no role in the design, conduct or analysis of the study nor in the interpretation and reporting of study findings.

Compliance with ethical standards Statements on human rights, the welfare of animals, informed consent, and the Declaration of Helsinki are not applicable to this manuscript. IRB approval is not applicable.

Ethical responsibilities of authors: All authors have approved the final manuscript and agree with its submission to Translational Behavioural Medicine. All authors have contributed equally to the scientific work and are responsible and accountable for the results. We confirm that this manuscript has not been previously published (partly or in full) and that the manuscript is not being simultaneously submitted elsewhere. We confirm that the data have not been previously reported elsewhere and that no data have been fabricated or manipulated to support our conclusions. No data, text or theories by others are presented as if they were the authors' own. The authors have full control of all data, which are accessible upon request.

Conflict of interest: OP, SM and AB declare that they have no conflict of interest. RW undertakes research and consultancy and receives fees for speaking from companies that develop and manufacture smoking cessation medications.

1. Rock Health. (2015). Digital health consumer adoption: 2015. Retrieved November 4, 2015, from http://rockhealth.com/reports/digital-health-consumer-adoption-2015/.
2. Fox S, Duggan M. Mobile health 2012. Pew Internet & American Life Project. 2012 Retrieved from http://www.pewinternet.org/2012/11/08/mobile-health-2012/.
3. Kontos, E., Blake, K. D., Chou, W.-Y. S., & Prestin, A. (2014). Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. *Journal of Medical Internet Research, 16*(7), e172. doi:10.2196/jmir.3117.
4. West, R., & Michie, S. (2016). *A Guide to Development and Evaluation of Digital Interventions in Healthcare*. London: Silverback Publishing.
5. Civljak, M., Stead, L. F., Sheikh, A., & Car, J. (2013). Internet-based interventions for smoking cessation. *Cochrane Database Syst Rev, 7*, CD007078.
6. Whittaker, R., Borland, R., Bullen, C., Rb, L., Mcrobbie, H., & Rodgers, A. (2009). Mobile phone-based interventions for smoking cessation. *Cochrane Database Syst Rev, 4*.
7. Nair, N. K., Newton, N. C., Shakeshaft, A., Wallace, P., & Teesson, M. (2015). A systematic review of digital and computer-based alcohol intervention programs in primary care. *Current Drug Abuse Reviews, 8*(2), 111–118.
8. Liu, F., Kong, X., Cao, J., Chen, S., Li, C., Huang, J., et al. (2015). Mobile phone intervention and weight loss among overweight and obese adults: a meta-analysis of randomized controlled trials. *Am J Epidemiol, 181*(5), 337–348. doi:10.1093/aje/kwu260.
9. Muntaner, A., Vidal-Conti, J., & Palou, P. (2015). Increasing physical activity through mobile device interventions: a systematic review. *Health Informatics Journal*, 1–19. doi:10.1177/1460458214567004.
10. Jones, K. R., Lekhak, N., & Kaewluang, N. (2014). Using mobile phones and short message service to deliver self-management interventions for chronic conditions: a meta-review. *Worldviews on Evidence-Based Nursing/Sigma Theta Tau International, Honor Society of Nursing, 11*(2), 81–88. doi:10.1111/wvn.12030.
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. *Journal of Medical Internet Research, 13*(3), e52. doi:10.2196/jmir.1772.
12. Cobb, N. K., Graham, A. L., Bock, B. C., Papandonatos, G., & Abrams, D. B. (2005). Initial evaluation of a real-world Internet smoking cessation system. *Nicotine & Tobacco Research, 7*(2), 207–216. doi:10.1080/14622200500055319.
13. Tate, D. F., Wing, R. R., & Winett, R. a. (2001). Using Internet technology to deliver a behavioral weight loss program. *J Am Med Assoc, 285*(9), 1172–1177. doi:10.1001/jama.285.9.1172.
14. Alexander, G. L., McClure, J. B., Calvi, J. H., Divine, G. W., Stopponi, M. A., Rolnick, S. J., et al. (2010). A randomized clinical trial evaluating online interventions to improve fruit and vegetable consumption. *Am J Public Health, 100*(2), 319–326. doi:10.2105/AJPH.2008.154468.
15. The Cochrane Collaboration. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. [Updated March 2011]*. (J. Higgins & S. Green, Eds.) 2011 Retrieved from www.cochrane-handbook.org.
16. Krishnan A What are academic disciplines? NCRM Working Paper Series*: ESRC National Centre for Research Methods* (2009).
17. Krishnan A *Five strategies for practising interdisciplinarity. NCRM Working Paper Series: ESRC National Centre for Research Methods*. (2009). Retrieved from http://eprints.ncrm.ac.uk/782/1/strategies_for_practising_interdisciplinarity.pdf.
18. Csikszentmihalyi, M. (1990). *Flow: the Psychology of Optimal Performance*. New York: Cambridge University Press.
19. Danaher, B. G., Boles, S. M., Akers, L., Gordon, J. S., & Severson, H. H. (2006). Defining participant exposure measures in web-based health behavior change programs. *Journal of Medical Internet Research, 8*(3), e15. doi:10.2196/jmir.8.3.e15.
20. Couper, M. P., Alexander, G. L., Zhang, N., Little, R. J. A., Maddy, N., Nowak, M. A., et al. (2010). Engagement and retention: measuring breadth and depth of participant use of an online intervention. *Journal of Medical Internet Research, 12*(4), e52. doi:10.2196/jmir.1430.
21. Eysenbach, G. (2005). The law of attrition. *Journal of Medical Internet Research, 7*(1), e11. doi:10.2196/jmir.7.1.e11.
22. Consumer Health Information Corporation. Motivating patients to use smartphone health apps. (2015). Retrieved August 10, 2015, from http://www.prweb.com/releases/2011/04/prweb5268884.htm.
23. Bennett, G. G., & Glasgow, R. E. (2009). The delivery of public health interventions via the Internet: actualizing their potential. *Annu Rev Public Health, 30*, 273–292. doi:10.1146/annurev.publhealth.031308.100235.
24. Brouwer, W., Oenema, A., Raat, H., Crutzen, R., De Nooijer, J., De Vries, N. K., & Brug, J. (2010). Characteristics of visitors and revisitors to an Internet-delivered computer-tailored lifestyle intervention implemented for use by the general public. *Health Educ Res, 25*(4), 585–595. doi:10.1093/her/cyp063.
25. Kelders, S. M., Kok, R. N., Ossebaard, H. C., & Van Gemert-Pijnen, J. E. W. C. (2012). Persuasive system design does matter: a systematic review of adherence to web-based interventions.

Journal of Medical Internet Research, 14(6), e152. doi:10.2196/jmir.2104.
26. Schubart, J. R., Stuckey, H. L., Ganeshamoorthy, A., & Sciamanna, C. N. (2011). Chronic health conditions and internet behavioral interventions: a review of factors to enhance user engagement. *Computers, Informatics, Nursing, 29*(2), 81–92. doi:10.1097/NCN.0b013e3182065eed.
27. Huberman, M. A., & Miles, M. B. (1994). Data management and analysis methods. In *Handbook of Qualitative Research* (pp. 428–443). Thousand Oaks: SAGE Publications.
28. O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *J Am Soc Inf Sci Technol, 59*(6), 938–955.
29. Ritterband, L. M., Thorndike, F. P., Cox, D. J., Kovatchev, B. P., & Gonder-Frederick, L. a. (2009). A behavior change model for internet interventions. *Ann Behav Med, 38*, 18–27. doi:10.1007/s12160-009-9133-4.
30. Short, C. E., Rebar, A. L., Plotnikoff, R. C., & Vandelanotte, C. (2015). Designing engaging online behaviour change interventions: a proposed model of user engagement. *The European Health Psychologist, 17*(1), 32–38.
31. Centre for Reviews and Dissemination, U. of Y. *Systematic reviews: CRD's guidance for undertaking reviews in healthcare.* (K. Khan, G. Ter Riet, J. Glanville, A. Sowden, & J. Kleijnen, Eds.) (2008). Retrieved from http://eprints.whiterose.ac.uk/1139/.
32. Dixon-Woods, M., Cavers, D., Agarwal, S., Annandale, E., Arthur, A., Harvey, J., et al. (2006). Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Med Res Methodol, 6*, 35. doi:10.1186/1471-2288-6-35.
33. Dixon-Woods, M., Bonas, S., & Booth, A. (2006). How can systematic reviews incorporate qualitative research? A critical perspective. *Qual Res, 6*(1), 27–44. doi:10.1177/1468794106058867.
34. Entwistle, V., Firnigl, D., Ryan, M., Francis, J., & Kinghorn, P. (2012). Which experiences of health care delivery matter to service users and why? A critical interpretive synthesis and conceptual map. *Journal of Health Services Research & Policy, 17*(2), 70–78. doi:10.1258/jhsrp.2011.011029.
35. Kazimierczak, K. A., Skea, Z. C., Dixon-Woods, M., Entwistle, V. A., Feldman-Stewart, D., N'Dow, J. M. O., & MacLennan, S. J. (2013). Provision of cancer information as a "support for navigating the knowledge landscape": findings from a critical interpretive literature synthesis. *Eur J Oncol Nurs, 17*(3), 360–369. doi:10.1016/j.ejon.2012.10.002.
36. Morrison, L., Yardley, L., Powell, J., & Michie, S. (2012). What design features are used in effective e-health interventions? A review using techniques from critical interpretive synthesis. *Telemedicine and e-Health, 18*(2), 137–144. doi:10.1089/tmj.2011.0062.
37. Anderson, P. J. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology, 8*(2), 71–82. doi:10.1076/chin.8.2.71.8724.
38. Thomson Reuters. EndNote X7. Philadelphia, USA 2013.
39. Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *J Clin Epidemiol, 46*(5), 423–429. doi:10.1016/0895-4356(93)90018-V.
40. Dixon-Woods, M., Sutton, A., Shaw, R., Miller, T., Smith, J., Young, B., et al. (2007). Appraising qualitative research for inclusion in systematic reviews: a quantitative and qualitative comparison of three methods. *Journal of Health Services Research & Policy, 12*(1), 42–47. doi:10.1258/135581907779497486.
41. Barbour, R. S. (2001). Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? *Br Med J, 322*, 1115–1117. doi:10.1136/bmj.322.7294.1115.
42. Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med, 6*(7), e1000097. doi:10.1371/journal.pmed.1000097.
43. Brown E, Cairns P. A grounded investigation of game immersion. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems.* (2004) 1297–1300. ACM. doi:10.1145/985921.986048.
44. Bianchi-Berthouze N, Kim WW, Patel D. Does body movement engage you more in digital game play? and why? In *Proceedings of the International Conference on Affective Computing and Intelligent Interaction.* 2007: 102–113.
45. Chou JC, Hung C, Hung Y. Design factors of mobile games for increasing gamers' flow experiences. In *Proceedings of the 2014 I.E. ICMIT.* 2014:137–139.
46. Sharek, D., & Wiebe, E. (2014). Measuring video game engagement through the cognitive and affective dimensions. *Simulation & Gaming, 45*, 569–592. doi:10.1177/1046878114554176.
47. Zhou, T. (2013). Understanding the effect of flow on user adoption of mobile games. *Personal & Ubiquitous Computing, 17*, 741–748. doi:10.1007/s00779-012-0613-3.
48. Oh, J., & Sundar, S. S. (2015). How does interactivity persuade? An experimental test of interactivity on cognitive absorption, elaboration, and attitudes. *J Commun, 65*, 213–236. doi:10.1111/jcom.12147.
49. Bouvier, P., Lavoue, E., & Sehaba, K. (2014). Defining engagement and characterizing engaged-behaviors in digital gaming. *Simulation & Gaming, 45*(4–5), 491–507. doi:10.1177/1046878114553571.
50. Schønau-Fog, H., & Bjørner, T. (2012). "Sure, I would like to continue": a method for mapping the experience of engagement in video games. *Bull Sci Technol Soc, 32*(5), 405–412. doi:10.1177/0270467612469068.
51. Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*(9), 641–661.
52. McClure, J. B., Shortreed, S. M., Bogart, A., Derry, H., Riggs, K., St John, J., et al. (2013). The effect of program design on engagement with an internet-based smoking intervention: randomized factorial trial. *Journal of Medical Internet Research, 15*(3), e69. doi:10.2196/jmir.2508.
53. Voils, C. I., King, H. A., Maciejewski, M. L., Allen, K. D., Yancy Jr., W. S., & Shaffer, J. A. (2014). Approaches for informing optimal dose of behavioral interventions. *Ann Behav Med, 48*, 392–401. doi:10.1007/s12160-014-9618-7.
54. Wang, J., Sereika, S. M., Chasens, E. R., Ewing, L. J., Matthews, J. T., & Burke, L. E. (2012). Effect of adherence to self-monitoring of diet and physical activity on weight loss in a technology-supported behavioral intervention. *Patient Preference and Adherence, 6*, 221–226.
55. Calleja, G. (2007). Digital game involvement. *Games & Culture, 2*(3), 236–260.
56. Lin, J. C.-C. (2007). Online stickiness: its antecedents and effect on purchasing intention. *Behav Inform Technol, 26*(6), 507–516. doi:10.1080/01449290600740843.
57. Han, J. Y., Kim, J.-H., Yoon, H. J., Shim, M., McTavish, F. M., & Gustafson, D. H. (2012). Social and psychological determinants of levels of engagement with an online breast cancer support group: posters, lurkers, and non-users. *J Health Commun, 17*(3), 356–371. doi:10.1080/10810730.2011.585696.Social.
58. Burns, C. G., & Fairclough, S. H. (2015). Use of auditory event-related potentials to measure immersion during a computer game. *Int J Hum Comput Stud, 73*, 107–114. doi:10.1016/j.ijhcs.2014.09.002.
59. Chiang, Y.-T., Lin, S. S. J., Cheng, C.-Y., & Liu, E. Z.-F. (2011). Exploring online game players' flow experiences and positive affect. *The Turkish Online Journal of Educational Technology, 10*(1), 106–114.
60. Chung, J., & Gardner, H. J. (2012). Temporal presence variation in immersive computer games. *International Journal of Human-Computer Interaction, 28*(8), 511–529. doi:10.1080/10447318.2011.627298.
61. Fang, X., Zhang, J., & Chan, S. S. (2013). Development of an instrument for studying flow in computer game play. *International Journal of Human-Computer Interaction, 29*(7), 456–470. doi:10.1080/10447318.2012.715991.
62. Harmat, L., Manzano, Ö. D., Theorell, T., Högman, L., Fischer, H., & Ullén, F. (2015). Physiological correlates of the flow experience during computer game playing. *Int J Psychophysiol, 97*, 1–7. doi:10.1016/j.ijpsycho.2015.05.001.
63. Hilvert-Bruce, Z., Rossouw, P. J., Wong, N., Sunderland, M., & Andrews, G. (2012). Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depressive disorders. *Behav Res Ther, 50*(7–8), 463–468. doi:10.1016/j.brat.2012.04.001.
64. Lefebvre, R. C., Tada, Y., Hilfiker, S. W., & Baur, C. (2010). The assessment of user engagement with eHealth content: the eHealth engagement scale. *J Comput-Mediat Commun, 15*, 666–681. doi:10.1111/j.1083-6101.2009.01514.x.
65. Martey, R. M., Kenski, K., Folkestad, J., Feldman, L., Gordis, E., Shaw, A., et al. (2014). Measuring game engagement: multiple methods and construct complexity. *Simulation & Gaming, 45*, 528–547. doi:10.1177/1046878114553575.
66. Morrison, L., Moss-Morris, R., Michie, S., & Yardley, L. (2014). Optimizing engagement with Internet-based health behaviour change interventions: comparison of self-assessment with and without tailored feedback using a mixed methods approach. *Br J Health Psychol, 19*, 839–855. doi:10.1111/bjhp.12083.
67. O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science & Technology, 61*(1), 50–69. doi:10.1002/asi.
68. Bossen, D., Buskermolen, M., Veenhof, C., de Bakker, D., & Dekker, J. (2013). Adherence to a web-based physical activity intervention for patients with knee and/or hip osteoarthritis: a mixed method study. *Journal of Medical Internet Research, 15*(10), e223. doi:10.2196/jmir.2742.
69. Geraghty, A. W. A., Torres, L. D., Leykin, Y. A. N., & Mun, R. F. (2012). Understanding attrition from international internet health interventions: a step towards global eHealth. *Health Promot Int, 28*(3), 442–452. doi:10.1093/heapro/das029.

70. Arden-Close, E. J., Smith, E., Bradbury, K., Morrison, L., Dennison, L., Michaelides, D., & Yardley, L. (2015). A visualization tool to analyse usage of web-based interventions: the example of positive online weight reduction (POWeR). *Journal of Medical Internet Research, 2*(1), e8. doi:10.2196/humanfactors.4310.
71. Carter, M. C., Burley, V. J., Nykjaer, C., & Cade, J. E. (2013). Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. *Journal of Medical Internet Research, 15*(4), e32. doi:10.2196/jmir.2283.
72. Chen, Z., Koh, P. W., Ritter, P. L., Lorig, K., Bantum, E. O. C., & Saria, S. (2015). Dissecting an online intervention for cancer survivors: four exploratory analyses of internet engagement and its effects on health status and health behaviors. *Health Educ Behav, 42*(1), 32–45. doi:10.1177/1090198114550822.
73. Christensen, H., Griffiths, K. M., & Farrer, L. (2009). Adherence in internet interventions for anxiety and depression. *Journal of Medical Internet Research, 11*(2), e13. doi:10.2196/jmir.1194.
74. Crutzen, R., Cyr, D., & de Vries, N. K. (2012). The role of user control in adherence to and knowledge gained from a website: randomized comparison between a tunneled version and a freedom-of-choice version. *Journal of Medical Internet Research, 14*(2), e45. doi:10.2196/jmir.1922.
75. Cussler, E. C., Teixeira, P. J., Going, S. B., Houtkooper, L. B., Metcalfe, L. L., Blew, R. M., et al. (2008). Maintenance of weight loss in overweight middle-aged women through the internet. *Obesity, 16*(5), 1052–1060. doi:10.1038/oby.2008.19.
76. Davies, C., Corry, K., Van Itallie, A., Vandelanotte, C., Caperchione, C., & Mummery, W. K. (2012). Prospective associations between intervention components and website engagement in a publicly available physical activity website: the case of 10,000 steps Australia. *Journal of Medical Internet Research, 14*(1), e4. doi:10.2196/jmir.1792.
77. Dennison, L., Morrison, L., Lloyd, S., Phillips, D., Stuart, B., Williams, S., et al. (2014). Does brief telephone support improve engagement with a web-based weight management intervention? Randomized controlled trial. *Journal of Medical Internet Research, 16*(3), e95. doi:10.2196/jmir.3199.
78. Glasgow, R. E., Christiansen, S. M., Kurz, D., King, D. K., Woolley, T., Faber, A. J., et al. (2011). Engagement in a diabetes self-management website: usage patterns and generalizability of program use. *Journal of Medical Internet Research, 13*(1), e9. doi:10.2196/jmir.1391.
79. Manwaring, J. L., Bryson, S. W., Goldschmidt, A. B., Winzelberg, A. J., Luce, K. H., Wilfley, D. E., & Taylor, C. B. (2008). Do adherence variables predict outcome in an online program for the prevention of eating disorders? *J Consult Clin Psychol, 76*(2), 341–346. doi:10.1037/0022-006X.76.2.341.
80. Morrison, C., & Doherty, G. (2014). Analyzing engagement in a web-based intervention platform through visualizing log-data. *Journal of Medical Internet Research, 16*(11), e252. doi:10.2196/jmir.3575.
81. Murray, E., White, I. R., Varagunam, M., Godfrey, C., Khadjesari, Z., & McCambridge, J. (2013). Attrition revisited: adherence and retention in a web-based alcohol trial. *Journal of Medical Internet Research, 15*(8), e162. doi:10.2196/jmir.2336.
82. Poirier, J., & Cobb, N. K. (2012). Social influence as a driver of engagement in a web-based health intervention. *Journal of Medical Internet Research, 14*(1), e36. doi:10.2196/jmir.1957.
83. Cugelman, B., Thelwall, M., & Dawes, P. (2011). Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. *Journal of Medical Internet Research, 13*(1), e17. doi:10.2196/jmir.1367.
84. Henshaw, H., McCormack, A., & Ferguson, M. A. (2015). Intrinsic and extrinsic motivation is associated with computer-based auditory training uptake, engagement, and adherence for people with hearing loss. *Front Psychol, 6*, 1–13. doi:10.3389/fpsyg.2015.01067.
85. Hsu, C.-L., & Lu, H.-P. (2004). Why do people play on-line games? An extended TAM with social influences and flow experience. *Information & Management, 41*, 853–868. doi:10.1016/j.im.2003.08.014.
86. McCabe, M. P., & Price, E. (2009). Attrition from an internet-based psychological intervention for erectile dysfunction: who is likely to drop out? *Journal of Sex & Marital Therapy, 35*(5), 391–401. doi:10.1080/00926230903065963.
87. Postel, M. G., de Haan, H. A., ter Huurne, E. D., van der Palen, J., Becker, E. S., & de Jong, C. A. J. (2011). Attrition in web-based treatment for problem drinkers. *Journal of Medical Internet Research, 13*(4), e117. doi:10.2196/jmir.1811.
88. Johansson, O., Michel, T., Andersson, G., & Paxling, B. (2015). Experiences of non-adherence to internet-delivered cognitive behavior therapy: a qualitative study. *Internet Interventions, 2*, 137–142. doi:10.1016/j.invent.2015.02.006.
89. Sainsbury, K., Mullan, B., & Sharpe, L. (2015). Dissemination of an online theory-based intervention to improve gluten-free diet adherence in coeliac disease: the relationship between acceptability, effectiveness, and attrition. *International Journal of Behavioral Medicine, 22*, 356–364. doi:10.1007/s12529-014-9416-4.
90. VanDeMark, N. R., Burrell, N. R., Lamendola, W. F., Hoich, C. A., Berg, N. P., & Medina, E. (2010). An exploratory study of engagement in a technology-supported substance abuse intervention. *Substance Abuse Treatment, Prevention, and Policy, 5*(10), 1–14.
91. Al-Asadi, A. M., Klein, B., & Meyer, D. (2014). Pretreatment attrition and formal withdrawal during treatment and their predictors: an exploratory study of the anxiety online data. *Journal of Medical Internet Research, 16*(6), e152. doi:10.2196/jmir.2989.
92. Habibović, M., Cuijpers, P., Alings, M., van der Voort, P., Theuns, D., Bouwels, L., et al. (2014). Attrition and adherence in a WEB-based distress management program for implantable cardioverter defibrillator patients (WEBCARE): randomized controlled trial. *Journal of Medical Internet Research, 16*(2), e52. doi:10.2196/jmir.2809.
93. 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.
94. Neve, M. J., Collins, C. E., & Morgan, P. J. (2010). Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial web-based weight loss program. *Journal of Medical Internet Research, 12*(4), e69. doi:10.2196/jmir.1640.
95. Nicholas, J., Proudfoot, J., Parker, G., Gillis, I., Burckhardt, R., Manicavasagar, V., & Smith, M. (2010). The ins and outs of an online bipolar education program: a study of program attrition. *Journal of Medical Internet Research, 12*(5), e57. doi:10.2196/jmir.1450.
96. Richardson, A., Graham, A. L., Cobb, N., Xiao, H., Mushro, A., Abrams, D., & Vallone, D. (2013). Engagement promotes abstinence in a web-based cessation intervention: cohort study. *Journal of Medical Internet Research, 15*(1), e14. doi:10.2196/jmir.2277.
97. Oinas-Kukkonen, H., & Harjumaa, M. (2009). Persuasive systems design: key issues, process model, and system features. *Commun Assoc Inf Syst, 24*(28), 486–501.
98. Hong, J.-C., Chiu, P.-Y., Shih, H.-F., & Lin, P.-S. (2012). Computer self-efficacy, competitive anxiety and flow state: escaping from firing online game. *The Turkish Online Journal of Educational Technology, 11*(3), 70–76.
99. Meischke, H., Lozano, P., Zhou, C., Garrison, M. M., & Christakis, D. (2011). Engagement in "my child's asthma", an interactive web-based pediatric asthma management intervention. *Int J Med Inform, 80*(11), 765–774. doi:10.1016/j.ijmedinf.2011.08.002.Engagement.
100. Boyle, E. A., Connolly, T. M., Hainey, T., & Boyle, J. M. (2012). Engagement in digital entertainment games: a systematic review. *Comput Hum Behav, 28*(3), 771–780. doi:10.1016/j.chb.2011.11.020.
101. Haines-Saah, R. J., Kelly, M. T., Oliffe, J. L., & Bottorff, J. L. (2015). Picture Me Smokefree: a qualitative study using social media and digital photography to engage young adults in tobacco reduction and cessation. *Journal of Medical Internet Research, 17*(1), e27. doi:10.2196/jmir.4061.
102. Kim, Y. H., Kim, D. J., & Wachter, K. (2013). A study of mobile user engagement (MoEN): engagement motivations, perceived value, satisfaction, and continued engagement intention. *Decis Support Syst, 56*, 361–370. doi:10.1016/j.dss.2013.07.002.
103. Ludden, G. D., van Rompay, T. J., Kelders, S. M., & van Gemert-Pijnen, J. E. (2015). How to increase reach and adherence of web-based interventions: a design research viewpoint. *Journal of Medical Internet Research, 17*(7), e172. doi:10.2196/jmir.4201.
104. Parks, A. C. (2014). A case for the advancement of the design and study of online positive psychological interventions. *J Posit Psychol, 9*(6), 502–508. doi:10.1080/17439760.2014.936969.
105. Horsch, C., Lancee, J., Beun, R. J., Neerincx, M. A., & Brinkman, W.-P. (2015). Adherence to technology-mediated insomnia treatment: a meta-analysis, interviews, and focus groups. *Journal of Medical Internet Research, 17*(9), e214. doi:10.2196/jmir.4115.
106. Funk, K. L., Stevens, V. J., Appel, L. J., Bauck, A., Brantley, P. J., Champagne, C. M., et al. (2010). Associations of internet website use with weight change in a long-term weight loss maintenance program. *Journal of Medical Internet Research, 12*(3), e29. doi:10.2196/jmir.1504.
107. Graham, A. L., Cha, S., Cobb, N. K., Fang, Y., Niaura, R. S., & Mushro, A. (2013). Impact of seasonality on recruitment, retention, adherence, and outcomes in a web-based smoking cessation intervention: randomized controlled trial. *Journal of Medical Internet Research, 15*(11), e249. doi:10.2196/jmir.2880.
108. Peels, D. A., Bolman, C., Golsteijn, R. H. J., De Vries, H., Mudde, A. N., van Stralen, M. M., & Lechner, L. (2012). Differences in reach and attrition between web-based and print-delivered tailored interventions among adults over 50 years of age: clustered

randomized trial. *Journal of Medical Internet Research, 14*(6), e179. doi:10.2196/jmir.2229.
109. Steinberg, D. M., Levine, E. L., Lane, I., Askew, S., Foley, P. B., Puleo, E., & Bennett, G. G. (2014). Adherence to self-monitoring via interactive voice response technology in an eHealth intervention targeting weight gain prevention among black women: randomized controlled trial. *Journal of Medical Internet Research, 16*(4), e114. doi:10.2196/jmir.2996.
110. Strecher, V. J., McClure, J., Alexander, G., Chakraborty, B., Nair, V., Konkel, J., et al. (2008). The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. *Journal of Medical Internet Research, 10*(5), e36. doi:10.2196/jmir.1002.
111. Wanner, M., Martin-Diener, E., Bauer, G., Braun-Fahrländer, C., & Martin, B. W. (2010). Comparison of trial participants and open access users of a web-based physical activity intervention regarding adherence, attrition, and repeated participation. *Journal of Medical Internet Research, 12*(1), e3. doi:10.2196/jmir.1361.
112. Jahangiry, L., Shojaeizadeh, D., Montazeri, A., & Najafi, M. (2014). Adherence and attrition in a web-based lifestyle intervention for people with metabolic syndrome. *Iranian Journal of Public Health, 43*(9), 1248–1258.
113. Kuijpers, W., Groen, W. G., Aaronson, N. K., & van Harten, W. H. (2013). A systematic review of web-based interventions for patient empowerment and physical activity in chronic diseases: relevance for cancer survivors. *Journal of Medical Internet Research, 15*(2), e37. doi:10.2196/jmir.2281.
114. Mahmassani, H. S., Chen, R. B., Huang, Y., Williams, D., & Contractor, N. (2010). Time to play? Activity engagement in multiplayer online role-playing games. *Transportation Research Record: Journal of the Transportation Research Board, 2157*, 129–137. doi:10.3141/2157-16.
115. Ferguson, M. A., & Henshaw, H. (2015). Computer and internet interventions to optimize listening and learning for people with hearing loss: accessibility, use, and adherence. *Am J Audiol, 24*, 338–343. doi:10.1044/2015.
116. Weston A, Morrison L, Yardley L, Van Kleek M, Weal M. Measurements of engagement in mobile behavioural interventions? In *Digital Health.* 2015:1–8.
117. Donovan, E., Mahapatra, P. D., Green, T. C., Chiauzzi, E., Mchugh, K., Hemm, A., et al. (2015). Efficacy of an online intervention to reduce alcohol-related risks among community college students. *Addiction Research & Theory, 23*(5), 437–447. doi:10.3109/16066359.2015.1043625.
118. Khadjesari, Z., Murray, E., Kalaitzaki, E., White, I. R., McCambridge, J., Thompson, S. G., et al. (2011). Impact and costs of incentives to reduce attrition in online trials: two randomized controlled trials. *Journal of Medical Internet Research, 13*(1), e26. doi:10.2196/jmir.1523.
119. An, L. C., Perry, C. L., Lein, E. B., Klatt, C., Farley, D. M., Bliss, R. L., et al. (2006). Strategies for increasing adherence to an online smoking cessation intervention for college students. *Nicotine & Tobacco Research, 8*(December), S7–S12. doi:10.1080/14622200601039881.
120. Brouwer, W., Kroeze, W., Crutzen, R., de Nooijer, J., de Vries, N. K., Brug, J., & Oenema, A. (2011). Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. *Journal of Medical Internet Research, 13*(1), e2. doi:10.2196/jmir.1639.
121. Cairns, P., Cox, A. L., Day, M., Martin, H., & Perryman, T. (2013). Who but not where: the effect of social play on immersion in digital games. *Int J Hum Comput Stud, 71*, 1069–1077. doi:10.1016/j.ijhcs.2013.08.015.
122. Morris, R. R., Schueller, S. M., & Picard, R. W. (2015). Efficacy of a web-based, crowdsourced peer-to-peer cognitive reappraisal platform for depression: randomized controlled trial. *Journal of Medical Internet Research, 17*(3), e72. doi:10.2196/jmir.4167.
123. Crutzen, R., Cyr, D., Larios, H., Ruiter, R. A. C., & De Vries, N. K. (2013). Social presence and use of internet-delivered interventions: a multi-method approach. *PLoS One, 8*(2), e57067. doi:10.1371/journal.pone.0057067.
124. Ben-Zeev, D., Kaiser, S. M., & Krzos, I. (2014). Remote "hovering" with individuals with psychotic disorders and substance use: feasibility, engagement, and therapeutic alliance with a text-messaging mobile interventionist. *Journal of Dual Diagnosis, 10*(4), 197–203. doi:10.1080/15504263.2014.962336.Remote.
125. Miller, A. S., Cafazzo, J. A., & Seto, E. (2014). A game plan: gamification design principles in mHealth applications for chronic disease management. *Health Informatics Journal*, 1–10. doi:10.1177/1460458214537511.
126. Brigham, T. J. (2015). An introduction to gamification: adding game elements for engagement. *Medical Reference Services Quarterly, 34*(4), 471–480. doi:10.1080/02763869.2015.1082385.
127. Richardson, C. R., Buis, L. R., Janney, A. W., Goodrich, D. E., Sen, A., Hess, M. L., et al. (2010). An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. *Journal of Medical Internet Research, 12*(4), e71. doi:10.2196/jmir.1338.
128. Leslie, E., Marshall, A. L., Owen, N., & Bauman, A. (2005). Engagement and retention of participants in a physical activity website. *Preventive, 40*, 54–59. doi:10.1016/j.ypmed.2004.05.002.
129. Irvine, A. B., Russell, H., Manocchia, M., Mino, D. E., Cox Glassen, T., Morgan, R., et al. (2015). Mobile-web app to self-manage low back pain: randomized controlled trial. *Journal of Medical Internet Research, 17*(1), e1. doi:10.2196/jmir.3130.
130. Lin, H., & Wu, X. (2014). Intervention strategies for improving patient adherence to follow-up in the era of mobile information technology: a systematic review and meta-analysis. *PLoS One, 9*(8), e104266. doi:10.1371/journal.pone.0104266.
131. Kok, G., Bockting, C., Burger, H., Smit, F., & Riper, H. (2014). Mobile cognitive therapy: adherence and acceptability of an online intervention in remitted recurrently depressed patients. *Internet Interventions, 1*, 65–73. doi:10.1016/j.invent.2014.05.002.
132. van den Berg, M. H., Ronday, H. K., Peeters, A. J., Voogt-van der Harst, E. M., Munneke, M., Breedveld, F. C., & Vliet Vlieland, T. P. M. (2007). Engagement and satisfaction with an internet-based physical activity intervention in patients with rheumatoid arthritis. *Rheumatology, 46*(3), 545–552. doi:10.1093/rheumatology/kel341.
133. Stark, S., Snetselaar, L., Piraino, B., Stone, A., Kim, S., Hall, B., & Burke, L. E. (2011). PDA self-monitoring adherence rates in two dialysis dietary intervention pilot studies: BalanceWise-HD and BalanceWise-PD. *J Ren Nutr, 21*(6), 492–498. doi:10.1053/j.jrn.2010.10.026.PDA.
134. Mohr, D. C., Duffecy, J., Ho, J., Kwasny, M., Cai, X., Burns, M. N., & Begale, M. (2013). A randomized controlled trial evaluating a manualized TeleCoaching protocol for improving adherence to a web-based intervention for the treatment of depression. *PLoS One, 8*(8), e70086. doi:10.1371/journal.pone.0070086.
135. Klein, M., Mogles, N., & Wissen, A. V. (2014). Intelligent mobile support for therapy adherence and behavior change. *J Biomed Inform, 51*, 137–151. doi:10.1016/j.jbi.2014.05.005.
136. McCambridge, J., Kalaitzaki, E., White, I. R., Khadjesari, Z., Murray, E., Linke, S., et al. (2011). Impact of length or relevance of questionnaires on attrition in online trials: randomized controlled trial. *Journal of Medical Internet Research, 13*(4), e96. doi:10.2196/jmir.1733.
137. Helander, E., Kaipainen, K., Korhonen, I., & Wansink, B. (2014). Factors related to sustained use of a free mobile app for dietary self-monitoring with photography and peer feedback: retrospective cohort study. *Journal of Medical Internet Research, 16*(4), e109. doi:10.2196/jmir.3084.
138. Whiteside, U., Lungu, A., Richards, J., Simon, G. E., Clingan, S., Siler, J., et al. (2014). Designing messaging to engage patients in an online suicide prevention intervention: survey results from patients with current suicidal ideation. *Journal of Medical Internet Research, 16*(2), e42. doi:10.2196/jmir.3173.
139. Jennings, M. (2000). Theory and models for creating engaging and immersive e-commerce websites. In *Proceedings of the 2000 ACM SIGCPR Conference on Computer Personnel Research* (pp. 77–85). New York: ACM.
140. Park, N., Min, K., Jin, S. A., & Kang, S. (2010). Effects of pre-game stories on feelings of presence and evaluation of computer games. *Int J Hum Comput Stud, 68*, 822–833. doi:10.1016/j.ijhcs.2010.07.002.
141. Hwang, M.-Y., Hong, J.-C., Hao, Y.-W., & Jong, J.-T. (2011). Elders' usability, dependability, and flow experiences on embodied interactive video games. *Educ Gerontol, 37*(8), 715–731. doi:10.1080/03601271003723636.
142. Chapman P, Selvarajah S, Webster J. Engagement in multimedia training systems. In *Proceedings of the 32nd Hawaii International Conference on System Sciences* 1999; 0: 1–9. Washington, DC: IEEE. doi:10.1109/HICSS.1999.772808.
143. Liu, S., Liao, H., & Pratt, J. A. (2009). Impact of media richness and flow on e-learning technology acceptance. *Comput Educ, 52*, 599–607. doi:10.1016/j.compedu.2008.11.002.
144. Miller, A. S., Cafazzo, J. A., & Seto, E. (2014). A game plan: gamification design principles in mHealth applications for chronic disease management. *Health Informatics Journal.* doi:10.1177/1460458214537511.
145. Lieberman, D. Z. (2006). Effects of a personified guide on adherence to an online program for alcohol abusers. *Cyberpsychology & Behavior, 9*(5), 603–607.
146. Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., et al. (2004). Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol, 23*(5), 443–451. doi:10.1037/0278-6133.23.5.443.
147. Borrelli, B. (2011). The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public*

Health Dent, 71, S52–S63. doi:10.1111/j.1752-7325.2011.00233.x.
148. Ubhi, H. K., Michie, S., Kotz, D., Wong, W. C., & West, R. (2015). A mobile app to aid smoking cessation: preliminary evaluation of SmokeFree28. *Journal of Medical Internet Research, 17*(1), e17. doi:10.2196/jmir.3479.
149. Chinn, P. L., & Kramer, M. K. (1991). *Theory and nursing: a systematic approach.* St. Louis: Mosby-Year Book.
150. Fiannaca A, La Rosa M, Rizzo R, Urso A, Gaglio S. An ontology design methodology for Knowledge-Based systems with application to bioinformatics. In *Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2012 I.E. Symposium.* 2012:85–91.
151. Weber, R. (2012). Evaluating and developing theories in the information systems discipline. *J Assoc Inf Syst, 13*(1), 1–30.
152. O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science & Technology, 61*(1), 50–69. doi:10.1002/asi.21229.
153. Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavorial medicine. *Ann Behav Med, 16*(3), 199–202.
154. Haukkala, A., Uutela, A., Vartiainen, E., Mcalister, A., & Knekt, P. (2000). Depression and smoking cessation: the role of motivation and self-efficacy. *Addict Behav, 25.* doi:10.1016/S0306-4603(98)00125-7.
155. Linde, J. A., Jeffery, R. W., Levy, R. L., Sherwood, N. E., Utter, J., Pronk, N. P., & Boyle, R. G. (2004). Binge eating disorder, weight control self-efficacy, and depression in overweight men and women. *Int J Obes, 28*(3), 418–425. doi:10.1038/sj.ijo.0802570.

Perski *et al. BMC Medical Informatics and Decision Making* (2017) 17:25
DOI 10.1186/s12911-017-0422-8

BMC Medical Informatics and Decision Making

RESEARCH ARTICLE Open Access

Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study

Olga Perski[1*], Ann Blandford[2], Harveen Kaur Ubhi[3], Robert West[3] and Susan Michie[1]

Abstract

Background: Public health organisations such as the National Health Service in the United Kingdom and the National Institutes of Health in the United States provide access to online libraries of publicly endorsed smartphone applications (apps); however, there is little evidence that users rely on this guidance. Rather, one of the most common methods of finding new apps is to search an online store. As hundreds of smoking cessation and alcohol-related apps are currently available on the market, smokers and drinkers must actively choose which app to download prior to engaging with it. The influences on this choice are yet to be identified. This study aimed to investigate 1) design features that shape users' choice of smoking cessation or alcohol reduction apps, and 2) design features judged to be important for engagement.

Methods: Adult smokers ($n = 10$) and drinkers ($n = 10$) interested in using an app to quit/cut down were asked to search an online store to identify and explore a smoking cessation or alcohol reduction app of their choice whilst thinking aloud. Semi-structured interview techniques were used to allow participants to elaborate on their statements. An interpretivist theoretical framework informed the analysis. Verbal reports were audio recorded, transcribed verbatim and analysed using inductive thematic analysis.

Results: Participants chose apps based on their immediate look and feel, quality as judged by others' ratings and brand recognition ('social proof'), and titles judged to be realistic and relevant. Monitoring and feedback, goal setting, rewards and prompts were identified as important for engagement, fostering motivation and autonomy. Tailoring of content, a non-judgmental communication style, privacy and accuracy were viewed as important for engagement, fostering a sense of personal relevance and trust. Sharing progress on social media and the use of craving management techniques in social settings were judged not to be engaging because of concerns about others' negative reactions.

Conclusions: Choice of a smoking cessation or alcohol reduction app may be influenced by its immediate look and feel, 'social proof' and titles that appear realistic. Design features that enhance motivation, autonomy, personal relevance and credibility may be important for engagement.

Keywords: Alcohol reduction, Behaviour change, Engagement, Excessive alcohol consumption, mHealth, Smartphone apps, Smoking cessation, Think aloud, Thematic analysis

* Correspondence: olga.perski.14@ucl.ac.uk
[1]Department of Clinical, Educational & Health Psychology, University College London, 1-19 Torrington Place, London WC1E 6BT, UK
Full list of author information is available at the end of the article

Background

Cigarette smoking and excessive alcohol consumption are two of the most serious global public health problems [1]. Behaviour change interventions delivered face-to-face by trained healthcare professionals have been developed to help tackle them [2, 3]. With technological developments, behavioural interventions can now be delivered remotely via digital platforms. Digital behaviour change interventions include any behaviour change programme delivered via websites, mobile phones, smartphone applications (apps) or wearables [4]. Smartphones are typically carried with the user throughout the day and can therefore facilitate the delivery of behavioural support "just-in-time", independent of geographical location [5–7]. Although only a minority of available smoking cessation and alcohol reduction apps have been rigorously evaluated in, for example, randomised controlled trials (RCTs), preliminary results suggest that apps might be effective in supporting smokers to quit and excessive drinkers to reduce their alcohol consumption [8–12]. In order to benefit from smoking cessation and alcohol reduction apps, users must identify and select which apps to download from the myriad available on the market [13, 14] and engage with them over time [15]. To our knowledge, no study has yet explored what factors are important in shaping this selection and subsequent engagement.

Although public health organisations such as the National Health Service in the United Kingdom (UK) and the National Institutes of Health in the United States (US) provide access to online libraries of publicly endorsed health apps (e.g. https://www.nhs.uk/oneyou/apps; https://www.nlm.nih.gov/mobile/) [16, 17], the majority of these accredited apps fail to act in accordance with data protection principles, such as encrypting personal information transmitted to developer or third-party servers [18]. There is also little evidence to suggest that users rely on these online libraries when searching for and selecting novel apps. Rather, the two most frequently used methods of identifying new apps are to search an online store and to seek recommendations from friends and family [19]. As there are currently more than 400 smoking cessation and 700 alcohol-related apps available on the market [13, 14], the onus is on the user to actively select which app to download. Notwithstanding a recent increase in the development and formal evaluation of theory- and evidence-informed apps within the research community [8–11, 20, 21], the majority of popular smoking cessation and alcohol reduction apps do not include 'behaviour change techniques' associated with higher quitting rates in face-to-face interventions and do not adhere to public health guidelines [13, 14, 22–26].

While popular smoking cessation and alcohol reduction apps vary in their specific approaches to behaviour change, commonalities in the techniques employed have been identified. For example, four independent content analyses of smoking cessation apps available in the US [13, 24], UK [23] and South Korean [26] versions of the iTunes Store/Google Play Store found that at least one of the following techniques was employed in a large proportion of the reviewed apps: self-monitoring (e.g. tracking cigarettes smoked or days smoke-free), feedback on progress, advising on how to quit, rewarding abstinence, supporting identity change and hypnosis [13, 23, 26]. Three independent content analyses of alcohol-related apps available in the US [25], Australian, [22] and UK [14] versions of the iTunes Store/Google Play Store found that although the majority of apps actively encouraged alcohol consumption, those promoting alcohol reduction commonly employed at least one of the following techniques: self-monitoring, feedback on progress (e.g. money saved from not buying alcohol), social support (e.g. dialling one's sponsor), psychoeducation (e.g. information on the negative effects of excessive alcohol use) and hypnosis (e.g. audio recordings to encourage relaxation) [14, 22, 25]. With regards to features aimed at promoting engagement, one review of smoking cessation apps found that tailoring of content was employed in 45% of apps [24] while another review identified a decline in the use of engagement features such as tailoring of content and rewards (e.g. points/badges) in smoking cessation apps between 2012 and 2014 (69.6% reducing to 45.3%) [23].

Due to the variable quality of available smoking cessation and alcohol reduction apps, an important goal is to determine how the design of evidence-based apps can be improved to attract users' attention in online stores and hence, increase their likelihood of being selected and engaged with [27]. The choice of any kind of app in an online store is likely to be influenced by visceral reactions to the app's design and affective responses to and cognitive processing of the app's known attributes [28–31]. Lasting positive first impressions of the visual appeal of websites are formed rapidly (within 50–500 milliseconds of exposure) and are primarily based on affective responses [28, 29]. While visual appeal was highlighted by users as important when choosing from pre-specified lists of apps (e.g. health apps, games for entertainment), factors such as perceived usefulness, personal relevance, positive user ratings and prior knowledge of brand names were also considered vital [30, 31]. There appears to be a lack of evidence as to how users freely choose smoking cessation and alcohol reduction apps in an online store and what factors shape their choice.

The potential benefits of apps depend not only on good choices by users but also on their subsequent engagement [15]. A positive association between engagement and intervention effectiveness has been observed

[32, 33], suggesting that the likelihood of successful behaviour change depends on engagement with the intervention [15, 34]. In the context of digital behaviour change interventions, engagement has been defined as 1) the extent of intervention use (i.e. amount, depth, duration, frequency), and 2) a subjective experience characterised by attention, interest and affect [35]. Although it is unclear what level of engagement is required for different kinds of digital behaviour change interventions to be effective, engagement with health apps has typically been found to be low; it has been estimated that 25% of such apps are not used more than once by each user and that less than 10% of users return seven days after their first use [36, 37]. It is therefore important to identify design factors that promote or detract from engagement with digital health products [38].

Although numerous intervention studies have identified demographic (e.g. age, gender, educational attainment) and psychological (e.g. motivation, mental health status) factors that predict engagement with digital behaviour change interventions, few studies have employed experimental designs to evaluate the effect of specific design features on engagement (see [35] for a systematic review of 117 articles). While evidence from RCTs indicates that features such as reminders and prompts [39], tailoring of content [40], contact with a healthcare professional [41] and simultaneous delivery of content (as opposed to sequential delivery) [42] positively influence engagement with computer- and web-delivered behaviour change interventions, little is known about the specific design features that influence engagement with smoking cessation and alcohol reduction apps.

Results from a secondary analysis of automatically recorded usage data from an RCT of a smoking cessation app indicated that users more frequently engaged with some tools compared with others (i.e. 'developing a quit plan', 'tracking smoking', 'viewing progress') [43]; however, the effect of particular design features (e.g. ease of use, tailoring of content, rewards) on engagement was not explored. In a formal consensus exercise, behaviour change and alcohol experts rated features such as ease of use, tailoring of content, feedback, aesthetic appeal and 'unique smartphone features' as likely to engage users with a novel alcohol reduction app [44]; however, it is unclear whether experts' views align with those of users from the target population. A cross-sectional survey of users' views on the functionality of an alcohol reduction app developed based on guidance from the National Institute of Clinical Excellence found that users largely held favourable views towards the app's features (e.g. an alcohol tracker, information on excessive alcohol use, notifications) [45]; however, users from the target population were not involved in the design of the app and survey respondents were not prompted to reflect on how the app's features might influence their engagement. A qualitative study that explored young adults' views on behaviour change apps and what factors contribute to their willingness to engage with such apps found that accuracy, security and immediate effects on mood were considered important for engagement while context-sensing software features and sharing on social media were considered off-putting [46]. However, no study to date has explored smokers' and drinkers' views on what design features are likely to be important for engagement with smoking cessation and alcohol reduction apps.

To better guide the selection of design features that can be included in future experimental studies (e.g. factorial RCTs), it would be useful to identify design features that smokers and drinkers judge to be important for engagement with smoking cessation and alcohol reduction apps. The present study therefore aimed to address the following two research questions through the use of qualitative methods:

1. What design features shape smokers' and drinkers' choice of smoking cessation and alcohol reduction apps?
2. What design features are judged by potential users to be important for engagement with smoking cessation and alcohol reduction apps?

Methods

Study design

The Consolidated Criteria for Reporting Qualitative Research checklist was used in the design and reporting of this study [47]. A think aloud methodology was used to address the first research question, which involved asking participants to verbalise their thoughts, impressions and feelings whilst engaging with an app of their choice [48]. The role of the researcher in a think aloud study is to retreat to the background and only prompt participants when necessary. This method was chosen over a retrospective design due to its ability to generate real-time data on the selections made, which was considered more reliable than data generated from participants' memory. Semi-structured interview techniques were used to allow participants to elaborate on statements made during the think aloud tasks and to address the second research question. Behaviour is often influenced by unconscious processing of stimuli [49], so users may have limited insight into the factors that in fact influence their engagement with apps. However, user-centred design methods emphasise the importance of exploring users' views as part of the iterative design process in order to develop digital behaviour change interventions that accommodate the needs of the target population [50–52].

Theoretical framework

As we were interested in exploring novel themes not previously identified in the literature, an interpretivist theoretical framework was used to inform data gathering and analysis [53]. Interpretivism proposes that multiple realities exist (i.e. assumes a 'subjective' rather than 'objective' reality) and that participants' accounts of their "lived experience" are co-constructed through the interaction with and subsequent interpretations of the researcher [53, 54]. Interpretivism recognises the active role of the researcher in both the elicitation and interpretation of qualitative data.

Participants

Smokers were eligible to take part if they i) were aged ≥ 18 years, ii) smoked cigarettes daily, iii) would consider using a smartphone app to help them stop smoking, iv) owned an Android or iOS smartphone with internet access that was capable of running apps and v) lived in or near London (UK). Drinkers were eligible to participate if they i) were aged ≥ 18 years, ii) reported an Alcohol Use Disorders Identification Test-Consumption (AUDIT-C) score ≥ 5, indicating excessive alcohol consumption [55], iii) would consider using a smartphone app to help them reduce their drinking, iv) owned an Android or iOS smartphone with internet access that was capable of running apps and v) lived in or near London (UK). Smokers and drinkers interested in using an app to stop or cut down were recruited in order to mimic real-world conditions and hence generate more valid data. It was expected that these participants would be able to more vividly imagine engaging with the apps compared with smokers and drinkers uninterested in using an app to stop or cut down [56]. For pragmatic reasons, no cut-off was imposed on cigarettes per day for including smokers in the study. As approximately 47% of English smokers are interested in using a digital intervention to stop [57], it was deemed more important to recruit smokers who were interested in using an app to stop rather than heavy or highly dependent smokers. Participants who were both smokers and drinkers were only asked about one kind of app; they were allowed to indicate a preference for what behaviour to focus on. Participants who had already tried to quit smoking/reduce their drinking using an app were not excluded. Participants who were not fluent English speakers were excluded.

Sampling

Participants were recruited through social media (e.g. Facebook, Twitter) and posters placed on central London university campuses. The recruitment materials stated that smokers and drinkers were invited to the laboratory to complete a few smartphone-based tasks and share their views on smoking cessation or alcohol reduction apps. Snowballing techniques were also used by asking participants to refer friends or family members interested in using an app to stop smoking or cut down on drinking [58]. Participants were recruited in batches of five until theoretical saturation was judged to have occurred (i.e. when no novel themes were identified) [59]. Preliminary data analysis was conducted after each batch of five participants to determine if more participants were needed.

Measures

Data were collected at baseline on: 1) age; 2) gender; 3) ethnicity, measured using the Office for National Statistics' index [60]; 4) socio-economic status, measured using the self-reported version of the National Statistics Socio-Economic Classification [61]; 5) nicotine dependence, measured using the Heaviness of Smoking Index (HSI) [62, 63]; a score ≥ 4 on the HSI indicates high nicotine dependence [63]; 6) patterns of alcohol consumption, measured using the Alcohol Use Disorder Identification Test-Consumption (AUDIT-C) [55, 64, 65]; an AUDIT-C score ≥ 5 indicates excessive alcohol consumption [55]; 7) motivation to stop smoking or cutting down on drinking, measured using the Motivation To Stop Scale (MTSS) [66]; 8) whether they had tried to stop/cut down in the past 12 months; 9) whether they had ever used an app to stop smoking/reduce drinking; 10) frequency of app use; 11) last time they had downloaded an app.

Procedure

Participants read the information sheet which described the nature of the study without disclosing information that might have influenced participants' search behaviours or verbal responses. They subsequently provided informed consent using an online screening questionnaire that assessed study eligibility and collected descriptive data (see Additional file 1). This questionnaire was hosted by Qualtrics survey software [67]. The face-to-face sessions were conducted in a private space at a London university or in participants' homes, according to participant preference. No one else was present besides the participant and researcher except for one interview that was conducted in a space where university students were present. Interviews took place between April and June 2016. Sessions lasted between 45 and 75 min. Participants received a £20 gift voucher as compensation for their time.

Pre-task interview

A pre-session interview was held to elicit participants' expectations of apps in general and smoking cessation or alcohol reduction apps in particular (see Additional file 2). Knowledge of participants' existing beliefs about apps and their smoking/drinking identity was judged to be relevant for the interpretation of subsequent statements and reactions; for example, knowledge that a participant did not identify as an excessive drinker was subsequently used to interpret ambiguous statements or reactions to the explored apps.

Think aloud tasks

Participants were instructed on how to think aloud (see Additional file 2) and were subsequently asked to complete a practice task: thinking aloud whilst changing the ringtone on their smartphone. Participants were then asked to complete two tasks on their smartphone. The first involved searching for smoking cessation or alcohol reduction apps in an online app store and was designed to elicit thoughts about factors that shape smokers' and drinkers' decisions to download such apps. The second task involved downloading and exploring a free smoking cessation or alcohol reduction app and was designed to gain insight into factors expected to be important for engagement (see Additional file 2). Positive reinforcement was used to ensure that participants verbalised relevant information (e.g. "You're doing well!"). When participants fell silent, prompts were used (e.g. "What are you thinking now?").

Debrief interview

The purpose of the debrief interview was to give participants the opportunity to elaborate on statements made during the think aloud tasks. Following the analysis of the first two batches of interview transcripts, the semi-structured interview schedule was adapted in order to elicit more data about points raised by the first 10 participants (see Additional file 2). At the end of the sessions, participants were told the full purpose of the study.

Data analysis

Sessions were audio-recorded, transcribed verbatim and analysed using inductive thematic analysis [68], which has previously been used to analyse data from think aloud studies involving smartphone apps [46, 69]. Braun and Clarke identify six phases of thematic analysis: i) familiarising with the data, ii) generating initial codes, iii) searching for themes, iv) reviewing themes, v) defining and naming themes, and vi) producing the report [68]. Data were coded by the first author using NVivo 10 [70] with regular discussions with the second author. New inductive codes were labelled as they were identified during the coding process. Data were sometimes assigned to multiple codes. All codes that potentially included data relating to the study aims were recorded. The first author reviewed the codes one by one, ordering the findings systematically under headings. The ordered data were reviewed and revised in discussion with the second author and were subsequently organised into themes. Theoretical saturation was judged to have occurred after 20 participants, as no new themes were identified [59]. As a quality check, the third author reviewed the codes, themes and participant quotes. Disagreements were resolved through discussion. Agreement on the final themes was reached through discussion between all co-authors. Differences between smokers and drinkers and other group differences were recorded where identified.

External validation

Respondent validation refers to the comparison of the researcher's interpretation of the data with participants' accounts to assess the level of agreement between the two [71, 72]. A subsample of five participants (25%) was contacted and asked to review the results after the initial themes had been developed. Participants were asked to comment on whether they felt that their views were well represented and the extent to which they agreed with the interpretation of their quotes and the main claims of the narrative. Three participants returned their comments, stating that they agreed with the authors' interpretations.

Reflexivity

Despite smoking and excessive drinking being associated with social stigma [73, 74], the interviewer felt that good rapport was built with the majority of participants. At the beginning of the study, the interviewer asked each participant the same set of questions in the same order, but it later became apparent that a more discursive style generated more extensive data and was therefore adopted.

Ethical approval

University College London's Departmental Research Ethics Committee granted ethical permission (UCLIC/1213/015). Personal identifiers were removed from the data, which were stored securely, and principles of research governance were observed [75].

Results

Participant characteristics

The average age of participants was 29.7 years ($SD = 9.2$), 60% were women, 70% were of White ethnicity, 20% were of Asian ethnicity, 85% were from a high socio-economic status background and 55% of participants had made an attempt to quit smoking or cut down on their drinking in the past 12 months but had relapsed into smoking/drinking (i.e. all participants were smoking/drinking at the time of the study). Smokers had an average HSI score of 0.6 ($SD = 1.07$), indicating low nicotine dependence, and drinkers had an average AUDIT-C score of 7.0 (*SD* = 2.9), indicating excessive alcohol consumption. Participant characteristics are found in Table 1.

Themes

Three themes were developed in relation to the first research question and were labelled "immediate look and feel of the app", "social proof" and "realistic and relevant titles". Five themes were developed in relation to the second research question and were labelled: "features that enhance motivation", "features that enhance autonomy", "features that enhance personal relevance", "features that enhance credibility" and "consistency with online and offline social preferences". As few differences between smokers and drinkers were identified, groups were combined for the reporting of the results unless otherwise stated. A summary of the identified themes is found in Table 2. Supplementary excerpts from the face-to-face sessions can be found in Additional file 3.

What factors shape smokers' and drinkers' choice of apps?

The immediate look and feel of the app

The majority of participants (14/20) stated that their choice of apps was guided by the initial appeal of icons and screenshots; however, the specific factors contributing to judgments about attractiveness differed across participants. Half of the participants (10/20) mentioned feeling drawn to apps using bright colours (e.g. light green, white), which were described as attention-grabbing or associated with health and wellbeing, while apps using dark or neon colours were considered less appealing. This divide was not universal; a few participants (2/20) felt more drawn to apps in dark colours because these were perceived as taking the quitting process more seriously.

> *Look at that! A dark screen, too many numbers. This really put me off. – D8*

When prompted to reflect on why particular designs caught their attention, many participants (9/20) mentioned that they preferred apps with minimalist or modern designs, as these were thought to signal professionalism and caring on the part of the developer, and described feeling "put off" by designs that looked "childish" or "amateurish". However, the majority of participants (11/20) were unable to articulate exactly what they liked about a particular design. This was manifested by statements about the app simply looking "nice" or having the "right" look.

Table 1 Participants' demographic, smoking, and drinking characteristics

ID	Group	Gender	Age	MTSS[a]	Made an attempt to stop/cut down in past 12 months	Ever used app to stop smoking or reduce drinking	Last time downloaded a smartphone app	Frequency of app use
D1	Drinker	M	24	5	Yes	No	In the last week	Daily
D2	Drinker	M	28	2	No	No	Today or yesterday	Daily
D3	Drinker	F	28	3	Yes	No	In the last month	Daily
D4	Drinker	F	31	6	No	No	In the last month	Weekly
D5	Drinker	F	21	2	No	No	Today or yesterday	Daily
D6	Drinker	F	56	2	No	No	In the last 6 months	Monthly
D7	Drinker	F	25	2	No	No	In the last 6 months	Daily
D8	Drinker	M	24	3	Yes	No	In the last month	Daily
D9	Drinker	M	47	3	Yes	No	In the last week	Daily
D10	Drinker	M	29	5	Yes	No	In the last week	Daily
S1	Smoker	M	24	2	No	No	In the last month	Several times/week
S2	Smoker	F	25	4	Yes	No	In the last week	Daily
S3	Smoker	M	28	3	No	No	In the last week	Daily
S4	Smoker	F	20	4	Yes	Yes	Today or yesterday	Daily
S5	Smoker	F	25	5	Yes	Yes	In the last week	Daily
S6	Smoker	F	27	7	Yes	No	In the last 3 months	Daily
S7	Smoker	M	25	2	No	No	In the last month	Daily
S8	Smoker	F	45	7	Yes	No	In the last 6 months	Daily
S9	Smoker	F	33	2	No	No	In the last week	Daily
S10	Smoker	F	28	5	Yes	No	In the last 3 months	Several times/week

[a]Motivation To Stop Scale (MTSS): 1 = I don't want to stop smoking/cut down on drinking alcohol, 2 = I think I should stop smoking/cut down on drinking alcohol but I don't really want to, 3 = I want to stop/cut down but haven't thought about when, 4 = I really want to stop/cut down but I don't know when I will, 5 = I want to stop/cut down and hope to soon, 6 = I really want to stop/cut down and intend to in the next 3 months, 7 = I really want to stop/cut down and intend to in the next month

Table 2 Summary of identified themes

	Theme	Description
1. What design features shape smokers' and drinkers' choice of apps?	The immediate look and feel of the app	First impressions of the app's aesthetic appeal (e.g. colour scheme, minimalist design) and usability (e.g. easy to understand, not too text-heavy).
	Social proof	The app's perceived quality, largely determined by 'social proof' (i.e. other users' ratings, recognition of credible brands/institutions).
	Realistic and relevant titles	Titles that appeared realistic and relevant to the target behaviour (e.g. "quit smoking", "reduce your drinking").
2. What design features are judged to be important for engagement?	Features that enhance motivation	Features that enhanced participants' motivation to stay smoke-free/reduce their drinking (e.g. monitoring and feedback, goal setting, rewards).
	Features that enhance autonomy	Features that enhanced participants' autonomy (e.g. user-controlled reminders, flexible quitting/reduction plans).
	Features that enhance personal relevance	Features that engendered a sense of personal relevance (e.g. tailoring of content, a non-judgmental communication style, gain-framed messages).
	Features that enhance credibility	Features that engendered a sense of credibility and trust (e.g. a clear privacy policy, information perceived to be accurate).
	Consistency with online and offline social preferences	Consistency with participants' attitudes towards sharing progress on social media or joining an online support community (i.e. online preferences) and their attitudes towards using the app to log cigarettes/units of alcohol or distract from cravings in social settings (i.e. offline preferences).

Don't like it, yeah. I can't say more, it's just intuitive, why. It's just not something I'd particularly want to look at. - S8

Many participants (9/20) mentioned that their choice was influenced by the app's perceived usability or simplicity, as they did not wish to invest time in apps that seemingly required too much effort, appeared to be overly complex or evoked confusion.

…they had these complicated graphs, and lots of information in your face, it would take you a while to read, whereas the app that I chose, it had information, it showed the progress, but it was much easier on the eye to read. - D1

Judgments about an app's ease of use were often interwoven with judgments about its aesthetic appeal (8/20), making it difficult to single out any one factor as being more important in guiding choice.

Social proof

The majority of participants (15/20) mentioned that taking other people's star ratings or reviews of apps into account was vital in guiding their choice due to the lack of other guidance as to which apps are of acceptable quality. Choosing a popular app over a less popular one, determined by their respective number of downloads or list position, was thought to save time due to not having to manually filter out poor quality apps.

…if an app has a good rating, despite the one or two people who are not satisfied, I think it would mean that it works for the majority of people. - S1

Many participants (8/20) mentioned feeling drawn to apps from familiar brands, organisations or developers; these were described as being more salient than other apps. When prompted to reflect on why they felt drawn to familiar brands, participants stated that they expected such apps to be of better quality than those from unknown brands; they were uninterested in information provided by developers or organisations lacking authority.

Who is […]? Whatever, I don't care, you know. It's just some guy who came up with an app. – S6

Realistic and relevant titles

Many participants (9/20) mentioned that the app's title was important in guiding their choice. Titles including key words such as "quit smoking" or "reduce your drinking" were considered appealing, as these appeared to provide a realistic summary of the app's content. Participants avoided apps with titles that sounded like advertisements, such as those including the word “now”. These were thought to make empty promises about being able to help participants without providing any evidence for their statements. A few drinkers (3/10) avoided titles including the word “alcoholic”, as they did not believe that such apps would be personally relevant.

I think the title is really, really important, in terms of, don’t give promises that... You’ve got to be really accurate and realistic, I think, to keep people interested. Don’t make claims like that, just easily. – S6

What factors are judged to be important for engagement?

Features that enhance motivation

The majority of participants (12/20) expected that regular monitoring of, for example, alcoholic beverages consumed or cigarettes smoked, and the receipt of feedback on their progress would be important for engagement. Being able to view a timeline of the days on which one had managed to stay smoke-free or drink less was expected to enhance motivation to continue, as participants did not want to “ruin their progress”.

That’s probably a big incentive to not smoke, because it’s just going to set that back to zero, and it’s showing you your ever increasing progress, so yeah, I do like that. - S4

Many participants (11/20) stated that they did not expect to re-engage with apps that were too difficult to use and/or confusing. A few participants (2/20) were particularly concerned that continuously opening the app to monitor their smoking or drinking would be too effortful and hence, lead to disengagement.

Many participants (8/20) mentioned that they expected goal setting to be engaging; they believed that the achievement of a goal would make them feel good about themselves and hence, increase their motivation to achieve further goals (i.e. a positive feedback loop).

If you set those manageable goals, so you could achieve it, if you feel like you’re actually progressing, getting something, then you’re more likely to go back. - D10

Of the 13 participants reacting to the provision of rewards within their selected app, approximately half (6/13) expected that the receipt of social or material rewards when achieving a goal, such as encouragement or badges, would increase their motivation to engage due to the desire to earn more rewards.

Doesn’t [the badge] motivate you to carry on? You want to get more to prove to yourself that you can get them. – D5

The other half of participants (7/13) were not convinced that earning virtual rewards would affect their motivation, as they did not attach any real value to intangible points or badges. A subtle difference between participants who had already tried to quit smoking or reduce their drinking in the past year and those who had not was observed; many (4/7) of those who had already tried to quit expressed negative attitudes towards the receipt of virtual rewards, perhaps suggesting that negative expectancy of such rewards might be linked to recent unsuccessful quit attempts.

I’m not really going to get any awards, am I? They’re not giving me any money or presents. - D8

Features that enhance autonomy

Of those expressing a desire to receive reminders to initiate engagement (11/20), the majority of these participants (9/11) wanted to control how frequently the app would contact them, as they had prior experiences of feeling bombarded or “bullied” by too many reminders.

…it was getting really, really annoying, and it bullied me a little bit too much, about me not meeting my goals that I set in the beginning when I started using it. Then it just went the other way, and it just went out the door, and I just took it off my phone. - S3

Many participants (9/20) already held firm beliefs about how to quit smoking or reduce their drinking. Smoking cessation apps that promoted a particular quitting strategy, such as quitting “cold turkey” with no option for gradual reduction, were therefore seen as inflexible. A few drinkers (4/10) expressed feeling annoyed with apps that rigidly compared their drinking patterns with the government’s recommended limits or persuaded users to have drink-free days, as they wanted to be in control of how to reduce their drinking in a meaningful way.

…it seems a bit extreme, especially when you’re not an alcoholic, why do you need a drink free day? Can’t you just have a small glass of wine with your meal? – D7

Features that enhance personal relevance

Tailoring of content according to individual preferences (13/20) inculcated a belief that the app was suited to the individual and that it was capable of providing effective support. For example, feedback on behavioural outcomes was estimated to be more engaging if it was tailored to the individual's needs and preferences.

I'm supposed to be motivated by how much money I've saved. That doesn't make sense to me. I think I should be motivated by how my health might have improved. I don't like this app. It's not going to help me. - D6

Information perceived as "preachy" or patronising made participants feel judged or nagged (9/20). This resulted in refusals to take the information seriously due to the desire to rebel against advice on what one "should" do.

I think I'm more likely to listen to practical advice rather than finger wagging… - S9

Some participants (6/20) mentioned that they wanted information about the positive effects of quitting or cutting down (i.e. 'gain-framed' messages). Information about health consequences that focused on the negative aspects of past smoking or drinking (i.e. 'loss-framed' messages) made participants (7/20) feel disempowered due to the inability to change past actions. Information focusing on the negative consequences of future smoking made some participants feel indifferent due to the inability to imagine one's future self.

Great. I started smoking when I was 13 and back then, I was smoking 40 cigarettes a day. - S3

A few drinkers (3/10) were sensitive to terminology perceived as "serious" or harsh, especially when terms such as "alcoholic" or "addict" were used. They were quick to distance themselves from apps using such terminology, as they appeared to assume that these must be catered to individuals who, unlike them, were dependent on alcohol. Smokers were more accepting of the use of the term "addict".

"Add an addiction." OK, quite serious… Wow! "I've been clean for…" That's some serious terminology. - D10

Features that enhance credibility

Many participants (8/20) mentioned that they felt uneasy about having to create an account with their personal e-mail address or allow access to the phone's location services in order to use their selected apps, as they were worried that their information would be passed on to third parties.

One thing is that I tend to not like apps that require so much data about my location services, because, I don't know, but obviously they sell on apps, so I think I'm quite wary of telling people too much about my data… - S10

However, a few participants (3/20) mentioned that their concerns were mitigated if a message about the app's policy on privacy and confidentiality was provided due to feelings of trust. A few participants (2/20) explicitly stated that they had no concerns regarding privacy in the context of apps.

It then says: "Your data will be anonymised and not shared with anybody other than for our research", which is nice to tell people for confidentiality reasons. - D7

Information judged to be inaccurate was met with scepticism by many participants (8/20) as errors and inconsistencies were thought to undermine the app's credibility. Participants did not want to waste time on inaccurate advice, as this was deemed to be untrustworthy.

I think it's really important that these sorts of sites and apps have the most current, up-to-date information, in order to get me to trust them, and take on board what they're telling me. - D2

Consistency with online and offline social preferences

Of the participants who reacted to the provision of social support features within their selected apps (10/20), such as sharing progress on social media (e.g. Facebook, Twitter) or joining an online community, few (4/10) expressed a desire to engage with such features; smoking and drinking were seen as private behaviours that are unacceptable to share with one's wider social network. Participants anticipated that sharing such information with others would generate pity rather than support.

…what do I want to get from that? I'm not going to get endorsements, I'm just going to get a few sad likes that are going to be quite patronising to me… - S3

A subtle difference was observed between those who had tried to quit smoking or reduce their drinking in the past year and those who had not; the former appeared to judge social sharing to not be engaging due to the anticipation of added pressure rather than increased support while the latter expressed more favourable attitudes towards social support features, especially those enabling users to join an online support community. Participants who had not made an attempt to quit expected that connecting with others in a similar situation might help stick to one's goals due to increased motivation.

Beliefs about the capability of apps to provide timely support when experiencing a craving were mixed. Many participants (7/20) struggled to see ways in which engagement with an app would influence their waning resolve. A few smokers (3/10) believed that doing a breathing exercise to assuage cravings would be helpful in the moment, but they did not want to use distraction games when socialising with others, who might find this behaviour strange.

> *Obviously, if you're in a bar, you're not going to be like: "I'm sorry guys, I just need to play my game." Maybe when you're home alone, it could be useful. – S5*

When imagining logging drinks consumed in social situations, a few drinkers (2/10) mentioned that they anticipated feeling embarrassed or uncomfortable, as others might find such behaviour "odd" or "rude" and hence, stop inviting them to the pub.

> *If I pull it out and start pressing it every time I've had a drink, they're going to start thinking that I'm odder than I really am. – D9*

Discussion

This study found that the immediate look and feel of apps, social proof and realistic and relevant titles shape smokers' and drinkers' choice of apps. Features that enhance motivation, including monitoring and feedback, goal setting, ease of use and rewards, and those that enhance autonomy, including flexible prompts and quitting strategies, were judged to be important for engagement. Participants also expected that features that engender a sense of personal relevance, such as tailoring of content according to individual preferences and the use of a non-judgmental communication style, and those that engender a sense of credibility, including privacy and accuracy, would be engaging. Moreover, consistency with one's online and offline social preferences was considered important for engagement. Few differences were found between smokers and drinkers.

The finding that the immediate look and feel of apps influenced participants' choice is consistent with the argument that visceral reactions to an app's design generate lasting positive first impressions [28, 29]. However, other people's app ratings and the perceived relevance of titles were also considered important. This supports the suggestion that both affective responses and cognitive processing of an app's attributes influence users' choice of apps [30, 31].

Our results are consistent with a number of well-established findings. Firstly, the finding that prompts, rewards, ease of use and tailoring of content according to individual differences were expected to be important for engagement supports previous research into computer-delivered smoking cessation and alcohol reduction interventions [42, 76, 77], results from content analyses of smoking cessation apps [23, 24] and findings from a formal expert consensus study [44]. Secondly, the finding that the app's communication style was judged to be important for engagement is consistent with previous research suggesting that the "tone of voice" of digital behaviour change interventions may evoke strong negative emotions and hence, cause participants to disengage [78]. Moreover, the finding that privacy and accuracy are expected to be important for engagement due to feelings of trust replicates research into other kinds of digital behaviour change interventions [79–81].

A frequently mentioned justification for using smartphone apps to deliver complex behaviour change interventions is that these are capable of delivering support as and when required, or "just-in-time" [5, 6]. As participants in the present study expressed concerns about engaging with smoking cessation and alcohol reduction apps in social settings due to anticipated embarrassment, this adds nuances to the assumption that smokers and drinkers want timely behavioural support irrespective of context. A recent study that employed geofencing (i.e. a software feature that uses the phone's global positioning system to set up geographical boundaries) to deliver context-aware smoking cessation support found that only a small proportion of pre-quit smoking reports (6.1%) were logged in social situations [82]. One of the reasons for this, as evidenced in follow-up interviews with participants, was fear of appearing rude to other people. This finding is also consistent with views expressed by young adults in a qualitative study exploring opportunities and challenges for behaviour change apps, who questioned the accuracy of context-sensing features [46].

Consistent with previous findings [46], smokers and drinkers in the present study did not want to share progress with their wider social networks due to the belief that others would pity rather than encourage this. It has been found that so-called 'closet' quit attempts (i.e. attempts to stop smoking without disclosure to anyone) are common among smokers [83]. As non-disclosure does not appear to be associated with a decreased likelihood of cessation success [83], this may be interpreted to suggest that social sharing should not be considered a 'one-size-fits-all' approach.

Care should be taken not to overstate the importance of the present findings due to the subtle group differences observed and the small sample size. However, it was found that attitudes towards joining an online support community and attitudes towards the receipt of virtual rewards appeared to differ depending on whether participants had made an attempt to quit/cut down in the past year. This suggests that individuals may differ in the factors that influence their judgments of engagement features. Future research should explore whether individuals may respond differently to social support features and rewards depending on their demographic and/or psychological characteristics.

Limitations

The method chosen to elicit data involved asking participants about their expectations about what factors would be engaging. As evidence suggests that the magnitude of relationships between beliefs and attitudes, intentions and actual behaviour are modest [84], further research is required to assess whether the inclusion of the features judged by participants to be important for engagement in the present study is in fact accompanied by higher levels of engagement. Although reliable methods for determining the potential of health apps to engage users (e.g. the Mobile Application Rating Scale [85]; a coding scheme developed by Ubhi and colleagues [86]) are available, the predictive validity of such scales (i.e. the scales' ability to predict actual levels of engagement) has not been evaluated. As the purpose of the present study was to explore smokers' and drinkers' views of apps, consistent with a user-centred approach to intervention design [50–52], think aloud methodology and semi-structured interview techniques were deemed to be more appropriate than existing quality scales. It has been argued that the use of think aloud methodology to elicit data might be problematic as it is cognitively demanding for participants to complete the assigned tasks whilst verbalising their thoughts [87]. However, we attempted to mitigate this issue by conducting debrief interviews to allow participants to elaborate on their statements.

The boundary between aesthetic appeal and perceived usability was often unclear in participants' explanations, highlighting the difficulty in articulating precisely why particular designs are considered more attractive than others and hence, indicating that the data generated here might be imprecise. However, ratings of beauty have been found to be strongly associated with ratings of perceived usability in other settings [88]. This emphasises the complexity of trying to dissociate these constructs and suggests that our findings are consistent with the published literature [28, 29]. Additional insight into how smokers and drinkers select apps (e.g. specific search terms used, non-conscious selection processes) might be gained from screen recordings or the use of eye tracking methodology.

As participants in the present study were predominantly of White ethnicity from high socio-economic status backgrounds and smokers indicated low levels of nicotine dependence it is possible that our findings do not generalise across the target population. However, participants reported similar levels of motivation to stop compared with a large, representative sample of English smokers ($N = 2483$): 35% in the present study versus 39% of English smokers in the earlier study indicated a MTSS score of ≥ 5 [66]. The finding that few smokers and none of the drinkers in the present study had ever used an app to quit smoking/reduce their alcohol consumption may be interpreted to suggest that the real concern is not how users decide which app to use, but rather, that it is more important to gain insight into what makes smokers and drinkers decide to use an app in the first place. Little is known about the uptake of smoking cessation and alcohol reduction apps in the general population of smokers and drinkers; however, findings from an ongoing series of cross-sectional household surveys of representative samples of the English population indicate that although half of smokers expressed an interest in using digital smoking cessation interventions (e.g. websites, smartphone apps), fewer than 1% had in fact used such interventions to support a quit attempt in the past year [57]. Hence, an alternative interpretation is that, according to available statistics, our sample appears similar to the target population with regards to previous app use.

Implications and future directions

As smokers and drinkers tend to select apps at least partly based on their immediate look and feel, it is important for healthcare professionals to collaborate with interaction design experts to develop evidence-based smoking cessation and alcohol reduction apps that are on a par with other commercially available apps in terms of aesthetics and usability. As participants in the present study were found to rely on 'social proof' (i.e. other users' ratings and brand recognition) when selecting apps, researchers and practitioners could leverage this by initiating collaborations with developers of popular apps or apps from well-known brands. For example, it might be more fruitful to modify the content of a well-established app with an existing client base rather than developing a novel smoking cessation or alcohol reduction app.

The finding that smokers and drinkers are more willing to engage with apps that provide options regarding quitting strategy poses a design challenge. As evidence suggests that some quitting strategies are more effective than others on average – for example, quitting smoking "cold turkey" is more effective than gradual reduction [89] – designers might benefit from using persuasive design elements, such as providing tutorials and guidance, using tunnelling techniques (i.e. making users click through a pre-specified sequence of pages), or making use of normative influence, to attempt to modify users' beliefs and attitudes [90].

Our findings suggest that the specifics of how to personalise content to support smokers' and drinkers' needs to promote engagement merit further investigation. A data-driven approach using machine-learning techniques

Table 3 Summary of design recommendations

Category	Design Recommendations
How can the reach of evidence-based apps be improved?	Develop smoking cessation and alcohol reduction apps that are on a par with other commercially available apps in terms of aesthetics and usability, perhaps through collaboration with interaction design experts.
	Researchers and practitioners may consider initiating collaborations with developers of popular apps and/or apps from well-known brands to leverage their existing 'social proof'.
	Use simple and straightforward titles that include key words (e.g. "quit smoking" or "reduce your drinking").
How can engagement be improved?	Use persuasive design elements (e.g. guidance, tunnelling, normative influence) to modify users' beliefs about how to quit smoking or reduce their drinking.
	Use machine-learning techniques to explore how to meaningfully tailor content according to individual differences (e.g. feedback, rewards).
	Develop response-sensitive notifications that tail off or adjust timings if the user stops reacting in order to prevent habituation or annoyance.
	Consider the online and offline social preferences of the target population. For example, it might be more fruitful to focus on action planning and/or behaviour substitution rather than in-the-moment support for smokers and drinkers.

might be helpful in advancing the knowledge on how to meaningfully tailor app content according to individual differences. For example, the type of feedback provided or whether or not to offer features that link users with others on social media could be tailored according to individual preferences to foster a sense of personal relevance. Furthermore, smokers and drinkers expected that too many reminders would lead to habituation and reduce autonomy. One means of preventing this is to develop response-sensitive notifications. For example, daily notifications could be sent as long as users react to these but their frequency would be reduced, or timing changed, as soon as users stop reacting to the prompts.

The finding that few smokers and drinkers wanted to use the apps in social settings suggests that the social context in which cigarette and alcohol cravings are triggered (e.g. pubs, cafés) should be considered in the design process. Smoking and drinking are perceived as more private than, for example, physical activity behaviours, perhaps due to social stigma [73, 74]. It should therefore not be assumed that features included in apps targeting other types of behaviour can successfully be transferred to those targeting smoking and drinking. The hypothesis that smokers and drinkers might engage more with apps that suggest how to replace smoking and drinking with other activities as opposed to those that provide in-the-moment support could be tested in future research. See Table 3 for a summary of design recommendations.

Conclusion

Smokers and drinkers interested in quitting or cutting down using a smartphone app choose apps based on their immediate look and feel, social proof and titles judged to be realistic and relevant. Features that enhance motivation, autonomy, personal relevance and credibility, and those that are consistent with users' online and offline social preferences are rated by participants as important for engagement.

Additional files

Additional file 1: Screening and baseline questionnaires. (DOCX 93 kb)

Additional file 2: Verbal instructions and semi-structured interview protocol. (DOCX 104 kb)

Additional file 3: Supplementary excerpts from the think aloud and interview sessions. (DOCX 110 kb)

Abbreviations

AUDIT-C: Alcohol use disorders test-consumption; HSI: Heaviness of smoking index; MTSS: Motivation to stop scale

Acknowledgements

We gratefully acknowledge all funding. We also acknowledge the members of UCL's Health Psychology Research Group for providing invaluable feedback on an early draft of this manuscript.

Funding

OP is a PhD candidate funded by a grant from Bupa under its partnership with University College London. SM is part-funded by grants from Cancer Research UK and the National Institute for Health Research School for Public Health Research. RW is funded by Cancer Research UK (C1417/A14135). The funders played no role in the design, conduct or analysis of the study, nor in the interpretation and reporting of study findings. The views expressed are those of the authors and not necessarily those of the funders.

Availability of data and materials

The full dataset supporting the conclusions of this article containing data not already included within the article or its additional files is available from the corresponding author on reasonable request.

Authors' contributions

OP, AB, RW and SM conceptualised the study design. OP recruited participants, conducted the think aloud sessions, analysed the data and drafted the first version of the manuscript. AB, HK, SM and RW provided guidance on the data collection, contributed to the data analysis and provided critical feedback on the manuscript. All authors approved the final manuscript.

Competing interests
OP, AB, HK and SM declare no competing interests. RW undertakes consultancy and research for and receives travel funds and hospitality from manufacturers of medications for smoking cessation.

Consent for publication
Consent for publication was obtained through the information sheet and consent form.

Ethics approval and consent to participate
Ethical approval was granted by UCL's Departmental Research Ethics Committee (UCLIC/1213/015). All participants read the information sheet and provided informed consent prior to taking part in this study.

Author details
[1]Department of Clinical, Educational & Health Psychology, University College London, 1-19 Torrington Place, London WC1E 6BT, UK. [2]UCL Interaction Centre, University College London, 66-72 Gower Street, London WC1E 6EA, UK. [3]Department of Epidemiology and Public Health, Cancer Research UK Health Behaviour Research Centre, University College London, 1-19 Torrington Place, London WC1E 6BT, UK.

Received: 26 October 2016 Accepted: 21 February 2017
Published online: 28 February 2017

References
1. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380:2224–60.
2. Kaner EFS, Dickinson HO, Beyer F, Pienaar E, Schlesinger C, Campbell F, et al. The effectiveness of brief alcohol interventions in primary care settings: A systematic review. Drug Alcohol Rev. 2009;28:301–23.
3. Lancaster T, Stead LF. Individual behavioural counselling for smoking cessation. Cochrane Database Syst Rev. 2005;(2):1–45.
4. West R, Michie S. A Guide to Development and Evaluation of Digital Interventions in Healthcare. London: Silverback Publishing; 2016.
5. Spruijt-Metz D, Nilsen W. Dynamic Models of Behavior for Just-in-Time Adaptive Interventions. IEEE Pervasive Comput. 2014;3:13–7.
6. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med. 2011;1:53–71.
7. Naughton F. Delivering "Just-In-Time" smoking cessation support via mobile phones: Current knowledge and future directions. Nicotine Tob. Res. 2016; Epub ahead of print.
8. Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend. 2014;143:87–94.
9. Buller DB, Borland R, Bettinghaus EP, Shane JH, Zimmerman DE. Randomized trial of a smartphone mobile application compared to text messaging to support smoking cessation. Telemed e-Health. 2014;20:206–14.
10. Ubhi HK, Michie S, Kotz D, Wong WC, West R. A Mobile App to Aid Smoking Cessation: Preliminary Evaluation of SmokeFree28. J Med Internet Res. 2015;17:e17.
11. Gustafson D, McTavish F, Chih M-Y, Atwood A, Johnson R, Boyle M. A smartphone application to support recovery from alcoholism: A randomized controlled trial. JAMA Psychiatry. 2014;71:566–72.
12. Gajecki M, Berman A, Sinadinovic K, Rosendahl I, Andersson C. Mobile phone brief intervention applications for risky alcohol use among university students: a randomized controlled study. Addict Sci Clin Pract. 2014;9:1–12.
13. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med. 2013;45:732–6.
14. Crane D, Garnett C, Brown J, West R, Michie S. Behavior change techniques in popular alcohol reduction apps: content analysis. J Med Internet Res. 2015;17:e118.
15. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. Am J Prev Med. 2016;51:833–42.
16. Public Health England. One You. 2016 [cited 2016 Oct 20]. Available from: https://www.nhs.uk/oneyou/apps
17. National Institues of Health. National Library of Medicine Mobile. 2016 [cited 2016 Oct 20]. Available from: https://www.nlm.nih.gov/mobile/
18. Huckvale K, Prieto JT, Tilney M, Benghozi P-J, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med. 2013;13:214.
19. The Nielsen Company. The Mobile Media Report - State Of The Media - Q3 2011. 2011.
20. Zhang MWB, Ward J, Ying JJB, Pan F, Ho RCM. The alcohol tracker application: an initial evaluation of user preferences. BMJ Innov. 2016;2:8–13.
21. Garnett C, Crane D, Michie S, West R, Brown J. Evaluating the effectiveness of a smartphone app to reduce excessive alcohol consumption: protocol for a factorial randomised control trial. BMC Public Health. 2016;16:536.
22. Weaver ER, Horyniak DR, Jenkinson R, Dietze P, Lim MSC. "Let's get Wasted!" and Other Apps: Characteristics, Acceptability, and Use of Alcohol-Related Smartphone Applications. J Med Internet Res. 2013;15:1–13.
23. Ubhi HK, Kotz D, Michie S, van Schayck OCP, Sheard D, Selladurai A, et al. Comparative analysis of smoking cessation smartphone applications available in 2012 versus 2014. Addict Behav. 2016;58:175–81.
24. Hoeppner BB, Hoeppner SS, Seaboyer L, Schick MR, Wu GWY, Bergman BG, et al. How Smart are Smartphone Apps for Smoking Cessation? A Content Analysis. Nicotine Tob Res. 2016;18:1025–31.
25. Cohn AM, Hunter-Reel D, Hagman BT, Mitchell J. Promoting Behavior Change from Alcohol Use through Mobile Technology: The Future of Ecological Momentary Assessment. Alcohol Clin Exp Res. 2011;35:2209–15.
26. Choi J, Noh GY, Park DJ. Smoking cessation apps for smartphones: Content analysis with the self-determination theory. J Med Internet Res. 2014;16:1–19.
27. Zhang MWB, Ho CSH, Cheok CCS, Ho RCM. Smartphone apps in mental healthcare: the state of the art and potential developments. BJPscyh Adv. 2015;21:354–8.
28. Lindgaard G, Fernandes G, Dudek C, Brown J. Attention web designers: You have 50 milliseconds to make a good first impression! Behav Inf Technol. 2006;25:115–26.
29. Norman DA. Emotional design: why we love (or hate) everyday things. New York: Basic Books; 2004.
30. Ayalew R. Consumer behaviour in Apple's App Store. Uppsala University; 2011.
31. Chang T-R, Kaasinen E, Kaipainen K. What Influences Users' Decisions to Take Apps into Use? A Framework for Evaluating Persuasive and Engaging Design in Mobile Apps for Well-Being. Proc. 11th Int. Conf. Mob. Ubiquitous Multimed. Ulm, Germany: ACM; 2012. p. 2.
32. Cobb NK, Graham AL, Bock BC, Papandonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. Nicotine Tob Res. 2005;7:207–16.
33. Tate DF, Wing RR, Winett RA. Using Internet technology to deliver a behavioral weight loss program. J. Am. Med. J Am Med Assoc. 2001;285:1172–7.
34. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res. 2011;13:e52.
35. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med. 2016;1–14.
36. Consumer Health Information Corporation. Motivating Patients to Use Smartphone Health Apps. 2015 [cited 2015 Aug 10]. Available from: http://www.prweb.com/releases/2011/04/prweb5268884.htm.
37. Appboy. Spring 2016 Mobile Customer Retention Report: An Analysis of Retention by Day. 2016.
38. Marcano Belisario J, Jamsek J, Huckvale K, O'Donoghue J, Morrison C, Car J. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. Cochrane Database Syst Rev. 2015;(7):1–113.
39. Lin H, Wu X. Intervention Strategies for Improving Patient Adherence to Follow-Up in the Era of Mobile Information Technology: A Systematic Review and Meta-Analysis. PLoS One. 2014;9:e104266.
40. Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in Web-based health behavior change programs. J Med Internet Res. 2006;8:e15.
41. Dennison L, Morrison L, Lloyd S, Phillips D, Stuart B, Williams S, et al. Does Brief Telephone Support Improve Engagement With a Web-Based Weight Management Intervention? Randomized Controlled Trial. J Med Internet Res. 2014;16:e95.

42. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res. 2008;10:e36.
43. Heffner JL, Vilardaga R, Mercer LD, Kientz JA, Bricker JB. Feature-level analysis of a novel smartphone application for smoking cessation. Am. J. Drug Alcohol Abuse. 2015;41:68–73.
44. Garnett C, Crane D, West R, Brown J, Michie S. Identification of Behavior Change Techniques and Engagement Strategies to Design a Smartphone App to Reduce Alcohol Consumption Using a Formal Consensus Method. JMIR mHealth uHealth. 2015;3:e73.
45. Zhang MWB, Fang P, Ho RCM. Global outreach and user preferences of a smartphone application developed for drinkers. Technol Heal Care. 2016;24:495–501.
46. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res. 2013;15:1–17.
47. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Heal Care. 2007;19:349–57.
48. Ericsson AK, Simon HA. How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. Mind Cult Act. 1998;5:178–86.
49. De Ridder D. Nudging for beginners: A shortlist of issues in urgent need of research. Eur Heal Psychol. 2014;16:2–6.
50. Maguire M. Methods to support human-centred design. Int J Hum Comput Stud. 2001;55:587–634.
51. Yardley L, Morrison LG, Andreou P, Joseph J, Little P. Understanding reactions to an internet-delivered health-care intervention: accommodating user preferences for information provision. BMC Med Informatics Decis. 2010;10:52.
52. Yardley L, Morrison L, Bradbury K, Muller I. The Person-Based Approach to Intervention Development: Application to Digital Health-Related Behavior Change Interventions. J Med Internet Res. 2015;17:e301.
53. Schwandt TA. Constructivist, Interpretivist Approaches to Human Inquiry. In: Denzin NK, Lincoln YS, editors. Handb. Qual. Res. London: SAGE Publications; 1994. p. 118–37.
54. Ponterotto JG. Qualitative Research in Counseling Psychology: A Primer on Research Paradigms and Philosophy of Science. J Couns Psychol. 2005;52:126–36.
55. Bradley KA, Debenedetti AF, Volk RJ, Williams EC, Frank D, Kivlahan DR. AUDIT-C as a brief screen for alcohol misuse in primary care. Alcohol Clin Exp Res. 2007;31:1208–17.
56. Rosson MB, Carroll JM. Usability Engineering: Scenario-Based Development of Human-Computer Interaction. San Fransisco: Morgan Kaufmann Publishers; 2002.
57. Brown J, Michie S, Raupach T, West R. Prevalence and characteristics of smokers interested in internet-based smoking cessation interventions: cross-sectional findings from a national household survey. J Med Internet Res. 2013;15:e50.
58. Biernacki P, Waldorf D. Snowball sampling: Problems and Techniques of Chain Referral Sampling. Sociol Methods Res. 1981;10:141–63.
59. Morse JM. Theoretical Saturation. In: Lewis-Beck MS, Bryman A, Futing Liao T, editors. SAGE Encycl. Soc. Sci. Res. Methods. Thousand Oaks: SAGE Publications, Inc; 2004.
60. Office for National Statistics. Guidance and Methodology: Ethnic Group, National Identity and Religion. 2016 [cited 2016 Mar 25]. Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/measuring-equality/equality/ethnic-nat-identity-religion/ethnic-group/index.html.
61. Chandola T, Jenkinson C. The new UK National Statistics Socio-Economic Classification (NS-SEC); investigating social class differences in self-reported health status. J Public Health Med. 2000;22:182–90.
62. Borland R, Yong HH, O'Connor RJ, Hyland A, Thompson ME. The reliability and predictive validity of the heaviness of smoking index and its two components: Findings from the International Tobacco Control Four Country Study. Nicotine Tob Res. 2010;12:45–50.
63. Heatherton TF, Kozlowski LT, Frecker RC, Rickert W, Robinson J. Measuring the heaviness of smoking: using self-reported time to the first cigarette of the day and number of cigarettes smoked per day. Br J Addict. 1989;84:791–9.
64. Bush KR, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT Alcohol Consumption Questions (AUDIT-C). Arch Intern Med. 1998;158:1789–95.
65. Frank D, DeBenedetti AF, Volk RJ, Williams EC, Kivlahan DR, Bradley KA. Effectiveness of the AUDIT-C as a screening test for alcohol misuse in three race/ethnic groups. J Gen Intern Med. 2008;23:781–7.
66. Kotz D, Brown J, West R. Predictive validity of the Motivation To Stop Scale (MTSS): A single-item measure of motivation to stop smoking. Drug Alcohol Depend. 2013;128:15–9.
67. Qualtrics. Qualtrics Survey Software [Internet]. Provo, Utah; 2016. Available from: http://www.qualtrics.com/.
68. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3:77–101.
69. Morrison L, Moss-Morris R, Michie S, Yardley L. Optimizing engagement with Internet-based health behaviour change interventions: Comparison of self-assessment with and without tailored feedback using a mixed methods approach. Br J Health Psychol. 2014;19:839–55.
70. QSR International. NVivo 10. 2012.
71. Lincoln YS, Guba EG. Naturalistic Inquiry. Newbury Park: SAGE Publications; 1985.
72. Mays N, Pope C. Assessing quality in qualitative research. Br Med J. 2000; 320:50–2.
73. Graham HM. Smoking, Stigma and Social Class. J Soc Polit. 2012;41:83–99.
74. Room R. Stigma, social inequality and alcohol and drug use. Drug Alcohol Rev. 2005;24:143–55.
75. Department of Health. Research Governance Framework for Health and Social Care. 2nd ed. 2005.
76. Donovan E, Mahapatra Das P, Green TC, Chiauzzi E, Mchugh K, Hemm A, et al. Efficacy of an online intervention to reduce alcohol-related risks among community college students. Addict Res Theory. 2015;23:437–47.
77. Haines-Saah RJ, Kelly MT, Oliffe JL, Bottorff JL. Picture Me Smokefree: a qualitative study using social media and digital photography to engage young adults in tobacco reduction and cessation. J Med Internet Res. 2015;17:e27.
78. Habibović M, Cuijpers P, Alings M, van der Voort P, Theuns D, Bouwels L, et al. Attrition and adherence in a WEB-Based Distress Management Program for Implantable Cardioverter defibrillator Patients (WEBCARE): randomized controlled trial. J Med Internet Res. 2014;16:e52.
79. Bossen D, Buskermolen M, Veenhof C, de Bakker D, Dekker J. Adherence to a web-based physical activity intervention for patients with knee and/or hip osteoarthritis: a mixed method study. J Med Internet Res. 2013;15:e223.
80. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res. 2009;11:e13.
81. Morrison L, Yardley L, Powell J, Michie S. What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. Telemed e-Health. 2012;18:137–44.
82. Naughton F, Hopewell S, Lathia N, Schalbroeck R, Brown C, Mascolo C, et al. A Context-Sensing Mobile Phone App (Q Sense) for Smoking Cessation: A Mixed-Methods Study. JMIR mHealth uHealth. 2016;4:e106.
83. Carpenter MJ, Sterba KR, Boatright AS, West R. "Closet" quit attempts: Prevalence, correlates and association with outcome. Addiction. 2011;106:2214–20.
84. Armitage CJ, Conner M. Efficacy of the Theory of Planned Behaviour: A meta-analytic review. Br J Soc Psychol. 2001;40:471–99.
85. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. JMIR mHealth uHealth. 2015;3:e27.
86. Ubhi HK, Michie S, Phil D, Kotz D, Van Schayck OCP, Selladurai A, et al. Characterising smoking cessation smartphone applications in terms of behaviour change techniques, engagement and ease-of-use features. Transl. Behav. Med. 2015;1–8.
87. Branch JL. Investigating the Information-Seeking Processes of Adolescents. Libr Inf Sci Res. 2000;22:371–92.
88. Hassenzahl M. The Interplay of Beauty, Goodness, and Usability in Interactive Products. Human-Computer Interact. 2004;19:319–49.
89. Cheong Y, Yong H, Borland R. Does how you quit affect success? A comparison between abrupt and gradual methods using data from the International Tobacco Control Policy Evaluation Study. Nicotine Tob Res. 2007;9:801–10.
90. Oinas-Kukkonen H, Harjumaa M. Persuasive Systems Design: Key Issues, Process Model, and System Features. Commun Assoc Inf Syst. 2009;24:486–501.

Original Research

DIGITAL HEALTH

Digital Health
Volume 4: 1–15
© The Author(s) 2018
DOI: 10.1177/2055207618785841
journals.sagepub.com/home/dhj
SAGE

Engagement features judged by excessive drinkers as most important to include in smartphone applications for alcohol reduction: A mixed-methods study

Olga Perski[1,4], Dario Baretta[2], Ann Blandford[3,4], Robert West[5] and Susan Michie[1]

Abstract

Objective: Engagement with smartphone applications (apps) for alcohol reduction is necessary for their effectiveness. This study explored (1) the features that are ranked as most important for engagement by excessive drinkers and (2) why particular features are judged to be more important for engagement than others.

Methods: Two studies were conducted in parallel. The first was a focus group study with adult excessive drinkers, interested in reducing alcohol consumption using an app ($n_{groups} = 3$). Participants individually ranked their top 10 features from a pre-specified list and subsequently discussed their rankings. The second was an online study with a new sample ($n = 132$). Rankings were analysed using the intraclass correlation coefficient (ICC) to assess the level of agreement between raters for each study. Qualitative data were analysed using inductive thematic analysis.

Results: There was low agreement between participants in their rankings, both in the focus groups (ICC = 0.15, 95% confidence interval (CI) = 0.03–0.38) and the online sample (ICC = 0.11, 95% CI = 0.06–0.23). 'Personalisation', 'control features' and 'interactive features' were most highly ranked in the focus groups. These were expected to elicit a sense of benefit and usefulness, adaptability, provide motivational support or spark users' interest. Results from the online study partly corroborated these findings.

Conclusion: There was little agreement between participants, but on average, the features judged to be most important for inclusion in smartphone apps for alcohol reduction were personalisation, interactive features and control features. Tailoring on users' underlying psychological needs may promote engagement with alcohol reduction apps.

Keywords

Alcohol reduction, behaviour change, digital health, engagement, mHealth, mixed-methods, smartphone apps, focus groups

Submission date: 28 November 2017; Acceptance date: 6 June 2018

Introduction

Approximately 43% of the world's adults consume alcohol regularly.[1] Excessive alcohol consumption is a risk factor for a wide range of physical (e.g. cirrhosis of the liver, cancer, stroke) and mental (e.g. depression, anxiety) conditions.[2–5] Interventions designed to reduce excessive alcohol consumption, delivered face-to-face by trained healthcare professionals, are available in many countries.[6–8] However, rising demand and pressures on national health budgets mean these

[1]Department of Clinical, Educational & Health Psychology, University College London, UK
[2]Department of Psychology, University of Milano-Bicocca, Italy
[3]UCL Interaction Centre, University College London, UK
[4]UCL Institute of Digital Health, University College London, UK
[5]Department of Behavioural Science and Health, University College London, UK

Corresponding author:
Olga Perski, Department of Clinical, Educational and Health Psychology, University College London, 1-19 Torrington Place, London WC1E 6BT, UK.
Email: olga.perski.14@ucl.ac.uk
Twitter: @OlgaPerski

services are limited and not meeting needs. With the advance of technology, behavioural support can be delivered digitally via websites, text messages or smartphone applications (apps). Smartphone apps support the delivery of behavioural support in real time,[9] and have the potential to reach a large proportion of drinkers at a low cost per additional user. However, to benefit from smartphone apps for alcohol reduction, drinkers must engage with them.[10] Although the precise nature of the relationship between engagement and intervention effectiveness is as yet unclear – particularly in the context of apps for alcohol reduction – low engagement with health apps is typically observed.[11–13] Although many users download and try health apps, engagement is typically not sustained for more than a few occasions.[12,13]

'Engagement' with an app can be defined as the extent to which those who have access to it use it (e.g. how often, for how long) and the manner in which they use it (e.g. attentively).[15] Whether a user engages with a given health app depends on its design (e.g. its content and how that content is delivered), the context in which it is used (e.g. who the users are, where and for what purpose they are using the app) and whether the app succeeds in changing particular 'mechanisms of action', such as users' attitudes towards the target behaviour, skills to perform or avoid the target behaviour, or motivation to change.[15] One plausible explanation as to why many users disengage from health apps is hence that these do not reflect users' needs, values and circumstances.[14]

The design of health apps is often driven by the possibility of using technology, and not because the target group has expressed a need for such technology.[14] The terms 'co-design' and 'user-centred design' are used to denote design processes in which potential users influence whether and how a design takes shape.[17] The user-centred design process typically involves several iteratively executed stages of development, including a needs and requirements analysis, prototyping (i.e. building an early version of the software) and usability testing.[18] Although few direct comparisons of health apps designed with and without user involvement have been made (but see DeSmet et al.[19] for a meta-analysis of serious games designed with and without user involvement), user-centred design activities may help clarify the needs and preferences that have to be met for a particular digital intervention to be engaged with by the target group.[14,20–22] Approaches to identifying user needs include contextual inquiry or ethnography, which can be used to identify the key issues faced by the target group, and qualitative interviews or focus groups, which can be used to identify potential users' goals, needs and ideas for design.[23] When an initial prototype has been developed, usability testing can shed light on how the app can be refined to better meet users' needs.

Several smartphone apps that target alcohol reduction in adult populations have recently been developed, with different degrees of user involvement and different approaches to gathering user data. To the authors' knowledge, the Location-Based Monitoring and Intervention System for Alcohol Use Disorders was one of the first smartphone apps designed to support adults who meet the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders criteria for alcohol use disorders (AUDs) and included educational materials, feedback on alcohol consumption, advice on problem solving and craving management strategies, location-triggered alerts and advice on behaviour substitution.[24,25] Users participating in a 6-week pilot study were asked to provide feedback on the app's functionality and usability at the end of the trial; however, it is unclear whether their feedback was used to refine the app. The Addiction-Comprehensive Health Enhancement Support System (A-CHESS) was designed to support adult patients leaving residential treatment for AUDs and included audio-guided relaxations, location-triggered alerts and a panic button that would alert two designated contacts.[26] Focus groups were conducted with patients, family members, criminal justice personnel and primary care physicians to gather user needs prior to the development of A-CHESS.[27] The PartyPlanner app was designed to support alcohol reduction in university students through behavioural simulation ahead of a drinking event, and the monitoring of and tailored feedback on individuals' estimated blood alcohol concentrations.[28] At the end of a randomised controlled trial (RCT) of the PartyPlanner app, participants were asked to rate the app's usability, suitability and the likelihood of recommending the app to a friend. The Alcohol Tracker app was designed to facilitate self-monitoring of alcohol consumption and included an alcohol diary, educational materials, goal setting and notifications.[29] Although survey respondents were invited to rate the app's perceived usefulness, the survey did not assess the app's usability or engagement potential. The 'CET' app was designed by Danish psychiatrists and psychologists to deliver cue exposure therapy to adults with AUDs.[30] User feedback on an initial version of the app was gathered through focus groups, and the app was refined accordingly prior to conducting an RCT. The Drink Less app was designed to support alcohol reduction in adults and included normative feedback, action planning, goal setting, feedback, monitoring, identity change and cognitive bias retraining.[31] Although users were not involved in the design of the app, a usability study was conducted to gather user feedback and the app was refined prior to evaluating its components in a factorial RCT.[32]

Although many existing alcohol-reduction apps have involved users in the design process, thus increasing their engagement potential, the benefits of such user-centred design activities may be limited by involving only a small number of potential users in the design process. Although this allows researchers and designers to gain an in-depth understanding of users' needs, insights from a small number of highly motivated participants who are willing to take part in design sessions may not generalise to other target users. For example, although community drug and alcohol service users were involved in the design of DIAMOND, a web-based alcohol intervention, few new patients recruited from the same service were willing to be randomised in a feasibility trial, mainly due to expressing a strong preference for face-to-face treatment.[33]

The present study used a mixed-methods approach, combining focus group methodology with an online study, to identify engagement features judged by excessive drinkers as most important to include in smartphone apps for alcohol reduction. We conducted in-depth focus group discussions with a small sample, in parallel with an online study with a larger sample of excessive drinkers, to address the following research questions:

1. What engagement features are ranked most highly by potential users of alcohol reduction apps?
2. What reasons do potential users give for judging particular features to be more important for engagement than others?

Methods

Study design

Two parallel studies were conducted. The first was a focus-group study and the second was an online study. As both methods have a number of well-known strengths and weaknesses, data sources were triangulated to address the same research questions.

Focus groups are useful for gaining an in-depth understanding of participants' experiences, beliefs and motivations, and are particularly suitable when the interaction between participants is expected to yield additional insight into the topic of interest.[34] Hearing about others' experiences and views may stimulate discussion and allow participants to elaborate on ideas mentioned by other group members.[35] However, a key weakness is that focus groups may inhibit the expression of controversial opinions due to social conformity, thus restricting the understanding of the diversity of users' needs and preferences.[35]

Research conducted online benefits from being able to reach larger, geographically diverse samples. Hence, results from online surveys are more likely to generalise to other members of the target population than findings from focus groups. Despite these strengths, online surveys that require cognitive effort may suffer from 'satisficing', where respondents simply provide a satisfactory answer or randomly choose from response options.[36,37]

Participants

1. Focus groups. Drinkers were eligible to participate in one of the focus groups if they (i) were aged ≥ 18 years, (ii) lived in or near London (United Kingdom; UK), (iii) reported an Alcohol Use Disorders Identification Test (AUDIT) score of ≥ 8, indicating excessive alcohol consumption,[38] (iv) owned an Android or iOS smartphone with internet access, (v) were interested in using a smartphone app to reduce their drinking and (vi) had previously used a health or fitness app. It was expected that participants with prior experience of using a health or fitness app would be able to more vividly imagine whether a particular feature would be important for engagement and hence generate more valid data.

Participants were recruited online through Gumtree (www.gumtree.com) and Call for Participants (www.callforparticipants.com) in addition to posters placed on central London university campuses. The recruitment materials stated that drinkers were invited to the laboratory to contribute to a focus group discussion with other participants about how to design engaging smartphone apps for alcohol reduction.

Of the 48 participants who completed the screening questionnaire, 29 were eligible to take part. In total, 13 participants did not respond to any further study communication. Six participants cancelled prior to taking part. One participant failed to arrive on time. In total, nine participants took part in one of three focus groups, with three participants in each group (see Figure 1). The average age of participants was 30.0 years ($SD = 10.1$), 77.8% were female and 66.7% had a non-manual occupation. Participants had an average AUDIT score of 13.6 ($SD = 3.1$), indicating excessive alcohol consumption (see Table 1).

2. Online sample. A new sample of drinkers were eligible to participate in the online study if they met the inclusion criteria outlined above, with the exception of (ii) and (vi). Instead, participants had to reside in the UK and did not need prior experience of using a health or fitness app. As we wanted to explore generalisability, we chose to be less restrictive in the online sample. Eligible participants who did not pass a multiple-choice attention check at the end of the ranking task (i.e. "What is a professional support feature?") were excluded from the analysis.

Participants were recruited online through Prolific Academic (www.prolific.ac). The recruitment materials

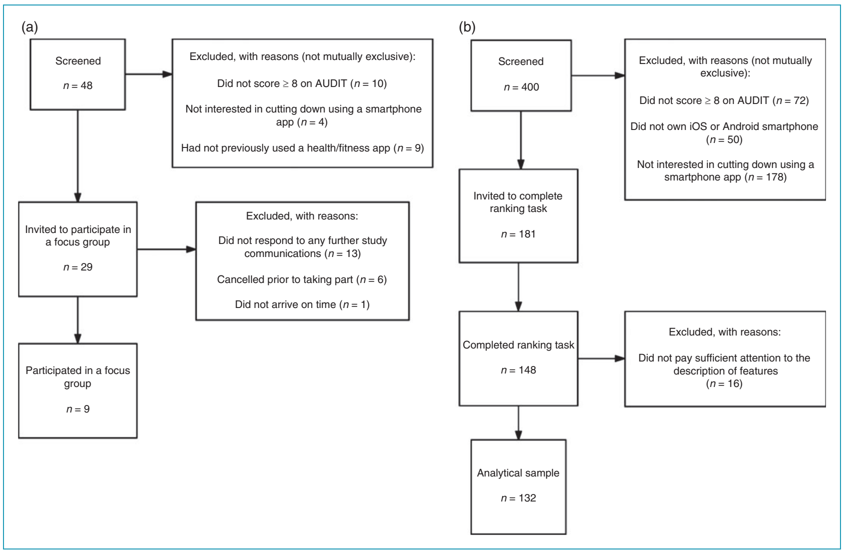


Figure 1. Participant flow charts for a) the focus group study, and b) the online sample.

invited drinkers to familiarise themselves with 16 different engagement features and rank their top 10 choices based on their likelihood of promoting engagement with apps for alcohol reduction.

Of 400 participants who completed the screening questionnaire, 181 were invited to complete the ranking task. Of these, 148 participants completed it, with 132 participants included in the analytical sample (see Figure 1). Just under half of the included participants were female (49.2%), 34.1% were aged 35–44 years, 13.6% had a manual occupation and 70.5% had a non-manual occupation. Participants had an average AUDIT score of 16.1 ($SD = 6.7$), indicating excessive alcohol consumption (see Table 1).

Measures

Data were collected on: (1) age; (2) gender; (3) occupational status (i.e. manual, non-manual, other); 4) alcohol consumption, measured using the AUDIT; (5) interest in using a smartphone app to help cut down on alcohol (yes vs. no); and (6) motivation to cut down on drinking alcohol, measured using the Motivation to Stop Scale (MTSS).

The AUDIT is a 10-item scale that taps three domains: alcohol consumption, drinking behaviour and alcohol-related problems. There is a maximum possible score of 40, with scores between 8 and 19 indicating excessive alcohol consumption, and scores of 20 or above indicating possible dependence.[38]

The MTSS is a single-item scale with seven response options: (1) I don't want to cut down on drinking alcohol; (2) I think I should cut down on drinking alcohol but I don't really want to; (3) I want to cut down but haven't thought about when; (4) I really want to cut down but I don't know when I will; (5) I want to cut down and hope to soon; (6) I really want to cut down and intend to in the next 3 months; (7) I really want to cut down and intend to in the next month. As the majority of available tools that tap motivation to reduce alcohol are based on the Stages of Change Model,[39] for which evidence is scarce,[40] the MTSS was used. Although the MTSS has yet only been validated in tobacco smokers,[41] it has been successfully employed in an observational study that estimated patterns of alcohol consumption and reduction in an English sample.[42]

Materials

In total, 16 different engagement features, derived from a relevant systematic review,[15] were used as stimuli (see Table 2). Feature descriptions were piloted and refined based on feedback from four independent researchers and five non-expert app users, recruited from the

Table 1. Participants' demographic and drinking characteristics.

Demographic and drinking characteristics	Focus groups, *n* (%)	Online sample, *n* (%)
Gender		
Women	7 (77.8%)	65 (49.2%)
Men	2 (22.2%)	67 (50.8%)
Age (years)		
18–24	4 (44.4%)	14 (10.6%)
25–34	3 (33.3%)	32 (24.2%)
35–44	0 (0%)	45 (34.1%)
45–54	2 (22.2%)	28 (21.2%)
55–64	0 (0%)	9 (6.8%)
65+	0 (0%)	4 (3.0%)
Occupational status		
Manual	0 (0%)	18 (13.6%)
Non-manual	6 (66.7%)	93 (70.5%)
Other	3 (33.3%)	21 (15.9%)
AUDIT, mean (*SD*)	13.6 (3.1)	16.1 (6.7)
MTSS		
1. I don't want to cut down on drinking alcohol	1 (11.1%)	8 (6.1%)
2. I think I should cut down on drinking alcohol but I don't really want to	1 (11.1%)	42 (31.8%)
3. I want to cut down but haven't thought about when	4 (44.4%)	16 (12.1%)
4. I really want to cut down but I don't know when I will	0 (0%)	10 (7.6%)
5. I want to cut down and hope to soon	1 (11.1%)	18 (13.6%)
6. I really want to cut down and intend to in the next 3 months	0 (0%)	10 (7.6%)
7. I really want to cut down and intend to in the next month	2 (22.2%)	28 (21.2%)

AUDIT = Alcohol Use Disorders Identification Test; MTSS = Motivation to Stop Scale.

authors' networks. Engagement features that have previously been found to be difficult for participants to describe verbally (e.g. aesthetics, ease of use, message tone) were not included. An experimental study design was expected to generate more valid data about how such abstract features influence engagement.[16]

Procedure

Interested participants read the information sheet describing the study. They subsequently provided informed consent via an online screening questionnaire, which also assessed study eligibility and collected descriptive data. The screening questionnaire was hosted by Qualtrics survey software.[43]

1. Focus groups. The focus groups were conducted at University College London. Sessions lasted approximately 2 hours. Participants received a £20 gift voucher as compensation for their time. Sessions were facilitated by the first author with support from the second author.

Table 2. Engagement features used in the ranking task.

Engagement features	Descriptions and examples
Challenge features	Features that allow you to compete against yourself or against other users, such as your friends. The app might, for example, encourage you to drink one unit fewer than your friends.
Control features	Features that allow you to make choices about how to use the app. The app might, for example, allow you to choose between a few different target goals instead of having one fixed option.
Action plans to use the app	A feature that encourages you to make a plan to use the app. An example might be to make a plan to open the app as soon as you have finished your breakfast every morning.
Setting a goal to use the app	A feature that encourages you to set a goal to use the app. For example, you might be able to set a goal to use the app once a day for two weeks.
Monitoring use of the app	A feature that allows you to record your use of the app. For example, the app might allow you to manually enter how much time you have spent on it, or it might record it automatically for you.
Feedback on use of the app	A feature that allows you to view your use of the app. For example, the app might show you how many times you have opened it on each day of the week.
Credibility features	Features that make you feel you can trust the app. For example, the app might have a clear privacy policy, be endorsed by a trusted organisation, or be free from adverts.
Guidance features	Features that explain how to use the app. This might, for example, include video tutorials about how the app works.
Interactive features	Features that allow, and respond to, input from the user. This might, for example, include a game or a knowledge quiz. The direct opposite would be a static app that does not allow you to enter any information or click into any of its features, much like this piece of text!
Novelty features	Features that ensure you see or learn something new every time you open the app. This might, for example, include daily content updates (e.g. a daily fact about alcohol or a daily motivational quote).
Narrative features	The presence of a storyline. For example, the app might be set up as a game or film with a plot, where you are the main character. This might include the presence of an avatar (i.e. a virtual figure that represents you).
Personalisation	Tailoring of content according to information about you (driven by the app) or customisation of the app so it looks or acts the way you prefer (driven by you). For example, the app might tailor its content based on information you give to it (e.g. about your age, gender, level of alcohol consumption) or you might be able to change the colour and font.
Professional support features	Features that enable you to have remote contact with a healthcare professional (e.g. the opportunity to chat to a nurse or a psychologist via the app).
Social support features	Features that allow you to connect with other app users. This might, for example, include an online discussion forum or a peer-to-peer instant messenger (e.g. a 'buddy' system).
Reminders to use the app	Regular push notifications or text messages that remind you to use the app.
Rewards for using the app	Being rewarded for using the app. You might, for example, receive a congratulatory message or a virtual badge/coin after having opened the app for seven days in a row.

Individual activity. An individual activity was first conducted to allow participants to familiarise themselves with the engagement features and elicit their attitudes to the features. The term 'engagement' was defined as a behaviour (e.g. how often you use the app, how much time you spend on it) and an experience (e.g. how interested you are in the app, how much attention you pay to it, how much you enjoy using it).[15]

Participants were each given a folder with Post-it Notes. Each of the 16 engagement features was described on a separate Post-it, accompanied by an illustrative example. Participants were also encouraged to think of their own examples. They were asked to rank their top 10 choices without consulting the other participants and were subsequently asked to place the Post-its with their selected features on a whiteboard, thus sharing their rankings with the group.

Group discussion. Participants subsequently convened to discuss their rankings. A semi-structured topic guide was used to steer the discussion (see Supplementary File 1). To gain a better understanding of why particular features were perceived as more important for engagement than others, participants were prompted to discuss the reasons for their rankings (e.g. "Can you tell me a bit more about why you ranked [insert feature here] highly?").

2. Online sample. Eligible participants were invited to complete the online ranking task in their own time on a personal computer, tablet or smartphone. The ranking task lasted for approximately 10 minutes and was hosted by Qualtrics survey software. Participants were paid £0.85 as compensation for their time. They were asked to complete the same ranking task as the focus group participants. At the end of the ranking task, participants were asked to respond to a multiple-choice attention check (described above). To gain a better understanding of why particular features were ranked more highly than others, participants were asked to respond to a free-text question about why they believed their top choice would be important for engagement.

Data analysis

1. Focus groups. Participants assigned a unique score from 1–10 to their top 10 engagement features, with 1 representing their top choice. The remaining six features were assigned a rank of 11, as the distance between these features was not expected to be meaningful. To assess the level of agreement between participants, the intraclass correlation coefficient (ICC) was estimated by means of a single measurement, absolute agreement, two-way, mixed-effects model. To assess whether some of the engagement features were, on average, ranked more highly than others, rankings were reverse scored (to aid interpretation) and descriptive statistics were calculated.

Sessions were audio recorded, transcribed verbatim and analysed using inductive thematic analysis. To inform the analysis, an interpretivist theoretical framework was used, based on the premise that the 'lived experience' of the individual can be captured through discussion between the researcher and participant.[44] The thematic analysis was conducted in six phases: (i) gaining familiarity with the data, (ii) generating initial codes, (iii) searching for themes, (iv) reviewing themes, (v) defining and naming themes and (vi) producing the report.[45] Data were coded independently by the first and second author. New inductive codes were labelled as they were identified during the coding process. Data were sometimes assigned to multiple codes. All codes that included data relating to the research questions were recorded. The first author reviewed the codes one by one, ordering the findings systematically under headings. The ordered data were reviewed and revised in discussion with the second author and were subsequently organised into themes. Disagreements were resolved through discussion. Agreement on the final themes was reached through discussion between all co-authors.

2. Online sample. Participants who provided incorrect responses to the 'attention check' were excluded from the analysis, as incorrect responses were interpreted to suggest that participants had not paid sufficient attention to the task to provide valid data.[37] A single measurement, absolute agreement, two-way, mixed-effects model was fitted to estimate the ICC. Rankings were reverse scored and descriptive statistics were calculated.

Responses to the free-text question about why participants believed their top choice would be important for engagement were analysed using inductive thematic analysis (described above).

Ethical approval

Ethical approval was granted by University College London's Departmental Research Ethics Committee (UCLIC/1213/015). Personal identifiers were removed and data were stored securely.

Results

1. Engagement features ranked most highly by potential users of alcohol reduction apps

1. Focus groups. There was positive but low agreement between participants (ICC = 0.15, 95% confidence interval (CI) = 0.03–0.38; see Figure 2). On average, participants ranked personalisation ($M = 8.67$, $SD = 2.12$), control features ($M = 7.22$, $SD = 3.73$) and interactive features ($M = 7.00$, $SD = 2.92$) most highly. Action plans ($M = 2.56$, $SD = 3.24$) and challenge features ($M = 2.67$, $SD = 2.40$) were judged to be the least important for engagement (see Table 3 and Figure 2).

2. Online sample. There was positive but low agreement between participants (ICC = 0.11, 95% CI = 0.06–0.23; see Figure 2). On average, participants ranked personalisation ($M = 6.74$, $SD = 3.18$), setting a goal to use the app ($M = 5.97$, $SD = 3.66$) and challenge features ($M = 5.56$, $SD = 3.93$) most highly. Narrative features ($M = 2.26$, $SD = 2.53$) and feedback on use of the app ($M = 2.68$, $SD = 2.33$) were judged to be least important for engagement (see Table 3 and Figure 2).

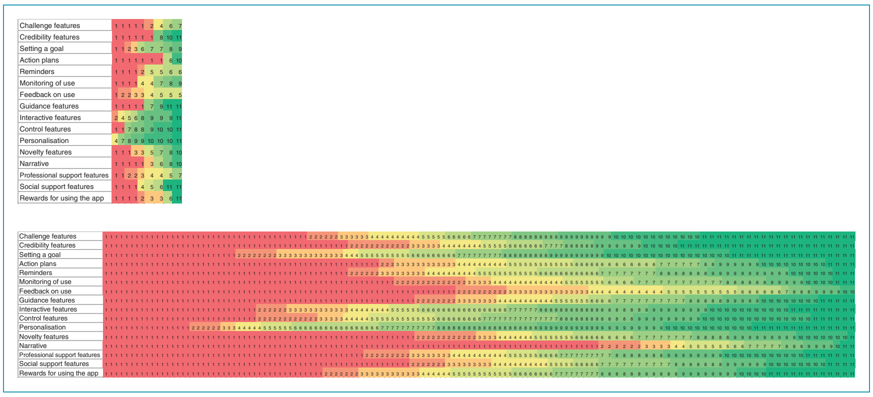


Figure 2. Heat maps of rankings in the focus groups (top), and in the online sample (bottom). Red, orange and yellow boxes indicate low rankings. Green boxes indicate high rankings.

Table 3. Mean rankings of the 16 engagement features in the a) focus groups ($n = 9$) and b) online sample ($n = 132$).

a) Focus groups		b) Online sample	
Engagement features	Mean (*SD*)	Engagement features	Mean (*SD*)
1. Personalisation	8.67 (*2.12*)	1. Personalisation	6.74 (*3.18*)
2. Control features	7.22 (*3.73*)	2. Setting a goal to use the app	5.97 (*3.66*)
3. Interactive features	7.00 (*2.92*)	3. Challenge features	5.56 (*3.93*)
4. Setting a goal to use the app	4.89 (*3.14*)	4. Interactive features	5.43 (*3.39*)
5. Guidance features	4.78 (*4.63*)	5. Control features	5.41 (*3.40*)
6. Social support features	4.56 (*4.13*)	6. Credibility features	4.86 (*3.99*)
7. Novelty features	4.33 (*3.35*)	7. Rewards for using the app	4.70 (*3.49*)
8. Monitoring of use	4.00 (*3.28*)	8. Professional support features	4.36 (*3.55*)
9. Credibility features	3.89 (*4.40*)	9. Reminders	4.27 (*3.20*)
10. Narrative features	3.56 (*3.54*)	10. Social support features	3.82 (*3.31*)
11. Feedback on use	3.33 (*1.50*)	11. Action plans	3.98 (*3.19*)
12. Professional support features	3.22 (*1.99*)	12. Guidance features	3.74 (*3.31*)
13. Rewards for using the app	3.22 (*3.35*)	13. Novelty features	3.66 (*3.16*)
14. Reminders	3.11 (*2.32*)	14. Monitoring of use	3.56 (*3.02*)
15. Challenge features	2.67 (*2.40*)	15. Feedback on use	2.68 (*2.33*)
16. Action plans	2.56 (*3.24*)	16. Narrative features	2.26 (*2.53*)

2. Judgments as to why particular features are expected to be more important for engagement than others

Six themes were generated: 'lack of trust and guidance as initial barriers', 'motivational support', 'benefit and usefulness', 'adaptability', 'sparking users' interest' and 'relatedness'. Two subthemes were developed in relation to the final theme, which were labelled 'perceived social stigma' and 'fear of social comparison' (see Table 4). Additional quotations can be found in Supplementary File 2.

1. Lack of trust and guidance as initial barriers. Although participants expected the presence of credibility features to be necessary to decide whether to engage with the app in the first place (as such features would inculcate feelings of trust), they did not believe that credibility features would promote further engagement after having made an initial decision to download an app.

> *...it wouldn't increase my engagement behaviour. It would just be the barrier, and make sure that I would actually use it, rather than frequently use it.* P2, focus group

Similarly, the presence of guidance features was expected to aid initial app navigation, but was not expected to prompt continued engagement. If guidance was provided again later, this was expected to be annoying, as participants believed they would be capable of using the app without any further support.

> *Just at the beginning of the app, when you've downloaded it and you're using it for the first time, it should tell you what to do. But not every time. You don't need guidance how to use it and where things are, because I think it would just be annoying.* P3, focus group

2. Motivational support. Participants expected features that provide motivational support to be important for engagement (e.g. control features, rewards, setting a goal to use the app, challenge features). This included

Table 4. Summary of themes and subthemes identified in a) the focus groups and b) the online sample.

Themes	Description	a) Identified in focus groups	b) Identified in online sample
1. Lack of trust and guidance as initial barriers	Features that inculcate feelings of trust and ensure the user can use the app comfortably (e.g. credibility features, guidance features) were considered more important for initial uptake than for continued engagement.	√	√
2. Motivational support	Features that support users' motivation to engage with the app or to cut down on drinking (e.g. control features, rewards, setting a goal to use the app, challenge features, message tone) were expected to encourage engagement, particularly if they promote users' independence.	√	√
3. Benefit and usefulness	Features that make users feel they are gaining something over and above *status quo* (e.g. personalisation, interactive features, novelty features, rewards) were expected to prompt engagement, particularly if they have utility 'in real life'.	√	√
4. Adaptability	Features that allow the app to adapt its content according to the user's level of progress or to intervene in the right moment (e.g. personalisation, interactive features, reminders) were expected to persuade the user and hence, promote engagement.	√	√
5. Sparking users' interest	Features that grab users' interest or provide a means of entertainment (e.g. narrative features, social support features, challenge features, interactive features, novelty features) were expected to prompt engagement.	√	√
6. Relatedness	Features that allow the user to connect with others who are in the same situation (e.g. social support features) were expected to promote engagement.	√	√
i. Perceived social stigma	Features that trigger app use in front of family and friends or connect users with close others (e.g. social support features, challenge features) were expected by some participants to elicit feelings of embarrassment and lead to disengagement.	√	
ii. Fear of social comparison	Features that encourage users to compete against friends or strangers (e.g. challenge features) were expected by some participants to be demoralising.	√	

features that support independent decision making by, for example, allowing users to make choices about how to use the app (e.g. control features). Participants expected to feel more motivated to work towards achieving goals they had set for themselves.

> *I feel that if you decide to carry out a task, you need to be in control of it, because ultimately, that's your goal that you're setting, and you want to have a sense of ownership or control of whatever you want to achieve. You feel more responsible for how you meet your goals.* P2, focus group
> *The more I would be able to manipulate the app to be and do what I wanted or needed, for my own circumstances, the more likely I am to use it.* P16, online sample

The app's 'tone of voice' or the way in which feedback was framed was expected to influence engagement. For example, feedback on drinking patterns framed in a positive manner (i.e. gain- rather than loss-framed) was expected to enhance users' beliefs about their ability to cut down on alcohol, and hence motivate engagement with the app.

> *...so that you don't feel discouraged when you drink too much, and then you decide that, you know what, I'm just going to ignore the app and shut it off.* P8, focus group

Participants believed that setting a goal to use the app or the receipt of rewards would motivate them to return to the app. For example, virtual rewards (e.g. badges, points) were expected to automatically encourage engagement.

> *It would encourage me to open the app on a daily basis.* P37, online sample
> *...even if it doesn't have practical meaning, it still works, because it's an incentive, and it tricks your brain to thinking that you're earning.* P3, focus group

Participants who ranked challenge features highly believed that competing against friends or other app users would help push oneself to achieve one's targets, thus providing an important source of motivation to cut down on drinking.

> *Personally, I feel if you have a community that challenges and pushes each other it encourages you to push yourself.* P47, online sample

3. Benefit and usefulness. Participants believed that features that make users feel they are gaining something over and above what they already knew or felt before downloading the app would be important for engagement (e.g. personalisation, interactive features, novelty features, rewards). For example, rewards that had utility 'in real life' or within the app itself (e.g. unlocking novel features, shopping vouchers) were thought to be more likely to prompt engagement due to their real-world usefulness.

> *Well, both of them are a kind of 'well done for doing this', they're both a reward, they both make you feel a bit better. But a badge, it's a cool fact, but it's not the same as having vouchers, where you can go and treat yourself to something you want.* P6, focus group

Maintaining a balance between the amount of effort on the part of the user (e.g. inputting vast amounts of information) and the rewards or outputs received from the app was expected to be crucial for engagement. Participants believed they would engage with the app only if they felt they were getting something meaningful back, such as learning something new about alcohol or about themselves (e.g. through personalised feedback). They also expected that they would feel more warmly towards apps that maintained a two-way flow of communication between user and app (i.e. 'reciprocal interactivity').

> *You've got to keep putting stuff in, but it's like, when am I going to get something out of it?* P5, focus group

Participants who did not rank narrative features, action plans or goal setting to use the app highly believed that such features would distract from the main task of reducing alcohol consumption or be more effortful than rewarding.

> *Well, surely the other features will make you want to use the app anyway.* P6, focus group

4. Adaptability. Participants expected features that make users feel that the app adapts itself to their level of progress or intervenes in the right moment (e.g. personalisation, interactive features, reminders) to promote engagement due to inculcating the belief that the app is speaking directly to the user. Highly personalised and context-sensitive information was expected to be more persuasive than generic advice about how to drink less.

> *If it's personal to me, you just get a sense of uniqueness, and you're like, yes, this is the best way for me to go, based on how I am right now.* P2, focus group
> *Every person is an individual, so I would have more faith in the app if it felt more tailored to my personal needs.* P34, online sample

Participants also expected features that allow the app to intervene either in the right moment or pre-emptively, 'before it is too late', would promote engagement. For example, participants who identified as heavy drinkers expected that professional support features would encourage engagement in 'times of crisis'.

> *It would help in times of crisis to be able to be in touch with a professional, or if I needed to ask health questions related to alcoholism.* P51, online sample

However, participants who did not identify as having a problem with alcohol did not expect professional support features to encourage engagement.

> *I think if I found that I had an issue with alcohol, maybe... – P9, focus group*

5. Sparking users' interest. Participants expected that the presence of features that grab users' attention or provide a means of entertainment (e.g. interactive features, narrative features, challenge features, social support features, novelty features) would prevent boredom and hence encourage users to return to the app. The hedonistic aspect of engagement was evident in participants' accounts, emphasising that some features are expected to be important for engagement only because they make the app more fun to use.

> *An app without any interactivity would get boring very quickly, and I would probably forget about it or delete it after a while.* P72, online sample
> *I do think that you need to keep people slightly entertained.* P9, focus group

Participants who ranked social support features highly believed that features that connect the user with others would draw their attention to the app and hence, promote engagement with other features.

> *If you saw a message from such and such, you might be more inclined to log on and respond to them. While you're on the app, you might use other features on it.* P6, focus group

6. Relatedness. Participants who ranked social support features highly expected that such features would facilitate the receipt of non-judgmental support from other users and hence, foster a sense of relatedness.

> *Being able to exchange feedback with strangers with the same goal could be supportive but non-judgemental as you will probably not know the other users.* P66, online sample .

i. Perceived social stigma. Participants who did not rank social support or challenge features highly imagined features that trigger app use in front of family or friends or connect users with others through the app would evoke feelings of embarrassment or worry that others may think they have a problem with alcohol.

> *I wouldn't want something like: 'Oh, why have you got that app?'* P5, focus group

ii. Fear of social comparison. Participants who did not rank social support or challenge features highly also pointed out that such features may have a negative effect on motivation to change due to eliciting fear of failure or worry that others are progressing quicker than oneself.

> *Somebody would always do better than me, performing better on the app than me, so I'd be engaging with people who are doing better than me on the app, which might be a bit demoralising.* P4, focus group

Discussion

Summary of main findings

This mixed-methods study found that there was low agreement between participants concerning the importance of particular engagement features, both in the focus groups and in the online sample. In general, features judged to be most important for inclusion in smartphone apps for alcohol reduction were personalisation, control features and interactive features. These features were expected to foster a sense of benefit and usefulness, adaptability, provide motivational support or spark users' interest. Social support and challenge features were ranked highly by a subset of participants as they were expected to foster relatedness and provide motivational support. However, another subset of potential users did not rank such features highly as they were expected to elicit social stigma or social comparison.

These findings lend support to and extend the results of prior research. First, there is previous support for the finding that personalisation is expected to promote engagement with alcohol reduction apps by inculcating the belief that the app is speaking directly to the user. Previous results have been consistent across types of study, including a formal expert consensus study[46] and a qualitative study with potential users.[16] This finding can be explained by the Elaboration Likelihood Model of Persuasion[47] and the Persuasive Systems Design Model,[48] which posit that messages tailored to users' needs and interests have greater potential for deep (as opposed to shallow) processing. Our findings highlight two additional mechanisms through which personalisation may promote engagement. First, personalisation may help to foster a sense of benefit and usefulness. For example, encouraging users to return to the app to learn more about themselves by offering highly personalised suggestions may prevent users from feeling that they are inputting data without getting anything back. Secondly, personalisation may help to foster a sense of adaptability by supporting both user-led and reactive use. For example, participants imagined they would engage more with apps that keep up-to-date with their progress and push relevant messages to users 'just-in-time'. Real-time message-tailoring based on current lapse risk has recently been deployed successfully in the smoking domain;[49] this strategy also merits investigation amongst excessive drinkers. Although existing apps for alcohol reduction have incorporated location-triggered alerts,[25,26] the utility of mood- or progress-triggered alerts is yet to be explored. A method that could be used to tailor messages in real-time is ecological momentary assessment, which has previously been used to assess drinking patterns and related cognitions and emotions.[50,51]

Secondly, previous research has emphasised the importance of features that support and develop users' motivation.[52–54] Participants in the present study highlighted that they would be more motivated to achieve goals they had set for themselves (i.e. 'autonomous motivation'), suggesting this kind of motivation may be more important for engagement than motivation that arises from external contingencies (i.e. 'controlled motivation').[55] However, the finding that participants also expected the receipt of rewards – which have previously been found to undermine autonomous motivation[56] – to help them engage, begs the question as to what sources of motivation are most supportive of engagement. This should be investigated experimentally (e.g. A/B testing or a factorial experiment). It may, for example, be hypothesised that features that support users' autonomous motivation will differentially impact on the total duration of engagement, as compared with features that support users' controlled motivation.

Thirdly, our results suggest that users may continue to engage with alcohol reduction apps only if they are regularly provided with information or features that pique their interest. Although few studies in the alcohol domain have highlighted the importance of preventing boredom, this is not a novel idea in the digital gaming and technology literature.[57,58] It has been argued that users have 'non-instrumental' needs (i.e. needs that do not serve as a means to achieve a particular aim), such as the need for stimulation or enjoyment.[59,60] The presence of features that address these non-instrumental needs is expected to give rise to a positive user experience and hence encourage technology engagement.[60] It has also been suggested that it may be particularly important to sustain users' interest in the technology when they have deviated from their goals.[61] The possibility of preventing disengagement due to relapse by providing features that meet users' need for stimulation should therefore be explored.

Fourthly, although findings from focus groups with young adults who drink at hazardous or harmful levels indicate a strong preference for features that foster relatedness,[62] evidence from studies with adult drinkers suggests that people typically react differently to features that connect them with friends or other users.[16] Our results suggest that excessive drinkers may either strongly like or dislike social support features or challenge features.

The finding that there were inconsistencies in participants' rankings begs the question as to how designers should prioritise features. By trying to satisfy everyone, we risk designing interventions that fit no one. However, as personalisation, interactive features and control features were generally preferred by excessive drinkers, a promising way forward may be to explore how these features could be embedded into alcohol reduction apps. It has been proposed that tailoring of content or features based on psychological constructs (e.g. the need for relatedness) is more effective than tailoring based on behaviour, which is in turn more effective than tailoring based on demographic characteristics.[63] Tailoring on users' underlying psychological needs, such as the need for relatedness, thus constitutes an important avenue for future research.

Limitations

This study was limited by employing an abstract, cognitively demanding ranking task that may have been more suitable for a face-to-face (as opposed to an online) study context. A plausible explanation as to why goal setting to use the app was ranked highly in the online sample is that users thought this referred to goal setting for alcohol reduction. We tried to limit misunderstandings by piloting the feature descriptions, but it is possible that some participants were still confused. Although participants' rankings should be interpreted with caution, the qualitative findings aid in the interpretation of the quantitative results.

It has been argued that users find it difficult to discuss design concepts without visual or tactile prompts, or that users are not designers.[64] Indeed, some participants in the present study found it difficult to articulate concrete design suggestions, such as how a narrative linked to alcohol reduction would pan out. However, as we did not want to limit participants' imagination of particular features, an abstract ranking task was deemed most suitable.

It is possible that the labels used for the engagement features may have biased participants' attitudes. This is suggested by a study in which old adults (aged 61-94 years) agreed that a 'falls-prevention intervention' was a good idea, but only for people who were older or frailer than them. The authors therefore concluded that reframing the intervention as a 'balance-training programme' might promote uptake.[65] In our study, labels such as 'professional support features' may have been perceived as too serious or irrelevant to participants' particular situations. This was suggested by a few participants. It is therefore possible that the finding that professional support features were preferred by participants who identified as being a 'heavy' drinker is an artefact of the labels used.

As men tend to exhibit more alcohol-related problems than women across countries,[66,67] the recruitment of more women than men into the focus groups constitutes a limitation. Future research should attempt to recruit a more balanced sample, with a view to exploring possible gender differences in app preferences. However, it should be noted that just over half of the online sample were male and we did not detect any differential preferences based on gender in this sample. Moreover, although the current approach to eliciting user needs provides useful information, an experimental study, in which the presence or design of particular features is manipulated, is required to test the actual impact on app engagement.

Conclusion

There was low agreement between participants concerning the importance of particular engagement features, but on average, those judged to be most important for inclusion in smartphone apps for alcohol reduction were personalisation, interactive features and control features. This study highlights that different features may be liked and used by different users, which should be considered in the design of novel alcohol reduction apps, or the modification of existing ones. Tailoring based on users' underlying psychological needs, such as the need for relatedness, constitutes an avenue for future research.

Acknowledgements: We gratefully acknowledge all funding. We also acknowledge the members of UCL's Health Psychology Research Group and Nikki Newhouse for providing feedback on an early draft of this manuscript.

Availability of data and materials: The full dataset supporting the conclusions of this article, containing data not already included within the article or its additional files, is available from the corresponding author on reasonable request.

Contributorship: OP, DB, AB, RW and SM conceived the study. OP, with support from DB, recruited participants, conducted the focus groups and analysed the data. OP drafted the first version of the manuscript. AB, RW and SM provided guidance on the data collection and contributed to the data analysis. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Declaration of conflicting interests: The author(s) declared the following potential conflicts of interest with respect to the research, authorship and/or publication of this article: OP, DB, AB and SM declare no competing interests. RW undertakes consultancy and research for and receives travel funds and hospitality from manufacturers of medications for smoking cessation.

Ethical approval: Ethical approval was granted by University College London's Departmental Research Ethics Committee (UCLIC/1213/015). All participants read the information sheet and provided informed consent prior to taking part in this study.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: OP is a PhD candidate funded by a grant from Bupa under its partnership with UCL. SM is part-funded by grants from Cancer Research UK and the National Institute for Health Research School for Public Health Research. RW is funded by Cancer Research UK (C1417/A22962). The funders played no role in the design, conduct or analysis of the study, nor in the interpretation and reporting of study findings. The views expressed are those of the authors and not necessarily those of the funders.

Guarantor: OP.

Peer review statement: This manuscript was reviewed by Vivian M. Gonzale, University of Alaska Anchorage, USA.

ORCID iD: Olga Perski http://orcid.org/0000-0003-3285-3174
Ann Blandford http://orcid.org/0000-0002-3198-7122

Supplementary Material: Supplementary material is available for this article online.

References

1. Gowing LR, Ali RL, Allsop S, et al. Global statistics on addictive behaviours: 2014 status report. *Addiction* 2015; 110: 904–919.
2. Boden JM and Fergusson DM. Alcohol and depression. *Addiction* 2011; 106: 906–914.
3. Dasarathy S and McCullough AJ. Alcoholic Liver Disease. In: Schiff ER, Maddrey WC and Sorrell MF (eds) *Schiff's Diseases of the Liver*. West Sussex, UK: John Wiley & Sons, Ltd., 2012, pp. 657–702.
4. Angus C, Holmes J, Pryce R, et al. *Alcohol and Cancer Trends: Intervention Scenarios*. 2016. University of Sheffield & Cancer Research UK.
5. Connor J. Alcohol consumption as a cause of cancer. *Addiction* 2016; 112: 222–228.
6. Kaner EFS, Dickinson HO, Beyer F, et al. The effectiveness of brief alcohol interventions in primary care settings: A systematic review. *Drug Alcohol Rev* 2009; 28: 301–323.
7. Jonas DE, Garbutt JC, Amick HR, et al. Behavioral counseling after screening for alcohol misuse in primary care: A systematic review and meta-analysis for the US Preventive Services Task Force. *Ann Intern Med* 2012; 157: 645–654.
8. Colom J, Scafato E, Segura L, et al. Brief interventions implementation on alcohol from the European health systems perspective. *Front Psychiatry* 2014; 5: 1–12.

9. Spruijt-Metz D and Nilsen W. Dynamic models of behavior for just-in-time adaptive interventions. *IEEE Pervasive Comput* 2014; 3: 13–17.
10. Yardley L, Spring BJ, Riper H, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med* 2016; 51: 833–842.
11. Eysenbach G. The law of attrition. *J Med Internet Res* 2005; 7: e11.
12. Appboy. *Spring 2016 Mobile Customer Retention Report: An Analysis of Retention by Day*. 2016.
13. Consumer Health Information Corporation. Motivating Patients to Use Smartphone Health Apps. http://www.prweb.com/releases/2011/04/prweb5268884.htm (2015, accessed 10 August 2015).
14. Kelders SM, Pots WTM, Oskam MJ, et al. Development of a web-based intervention for the indicated prevention of depression. *BMC Med Informatics Decis* 2013; 13: 1–11.
15. Perski O, Blandford A, West R, et al. Conceptualising engagement with digital behaviour change interventions: A systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2017; 7: 254–267.
16. Perski O, Blandford A, Ubhi HK, et al. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: A think aloud and interview study. *BMC Med Inform Decis Mak* 2017; 17: 1–14.
17. Vines J, Clarke R, Wright P, et al. Configuring participation: on how we involve people in design. In: *CHI '13: Changing Perspectives*. 2013, pp. 429–438.
18. International Standardisation Organisation. *ISO 9241-210. Ergonomics of Human-Computer Interaction - Part 2010: Human-Centred Design for Interactive Systems*. 2010.
19. DeSmet A, Thompson D, Baranowski T, et al. Is participatory design associated with the effectiveness of serious digital games for healthy lifestyle promotion? A meta-analysis. *J Med Internet Res* 2016; 18: e94.
20. Band R, Bradbury K, Morton K, et al. Intervention planning for a digital intervention for self-management of hypertension: A theory-, evidence- and person-based approach. *Implement Sci* 2017; 12: 1–13.
21. Yardley L, Ainsworth B, Arden-Close E, et al. The person-based approach to enhancing the acceptability and feasibility of interventions. *Pilot Feasibility Stud* 2015; 37: 1–7.
22. DeSmet A, Palmeira A, Beltran A, et al. The yin and yang of formative research in designing serious (exer-)games. *Games Health J* 2015; 4: 1–4.
23. Blandford A, Furniss D and Makri S. Qualitative HCI research: Going behind the scenes. *Synth Lect Human-Centered Informatics* 2016; 9: 1–115.
24. Giroux D, Bacon S, King DK, et al. Examining perceptions of a smartphone-based intervention system for alcohol use disorders. *Telemed e-Health* 2014; 20: 923–929.
25. Dulin PL, Gonzalez VM and Campbell K. Results of a pilot test of a self-administered smartphone-based treatment system for alcohol use disorders: Usability and early outcomes. *Subst Abus* 2014; 35: 168–175.
26. Gustafson D, McTavish F, Chih M-Y, et al. A smartphone application to support recovery from alcoholism: A randomized controlled trial. *JAMA Psychiatry* 2014; 71: 566–572.
27. Gustafson DH, Shaw BR, Isham A, et al. Explicating an evidence-based, theoretically informed, mobile technology-based system to improve outcomes for people in recovery for alcohol dependence. *Subst Use Misuse* 2011; 46: 96–111.
28. Gajecki M, Berman A, Sinadinovic K, et al. Mobile phone brief intervention applications for risky alcohol use among university students: A randomized controlled study. *Addict Sci Clin Pract* 2014; 9: 1–12.
29. Zhang MWB, Ward J, Ying JJB, et al. The alcohol tracker application: An initial evaluation of user preferences. *BMJ Innov* 2016; 2: 8–13.
30. Mellentin AI, Stenager E, Nielsen B, et al. A smarter pathway for delivering cue exposure therapy? The design and development of a smartphone app targeting alcohol use disorder. *JMIR mHealth uHealth* 2017; 5: e5.
31. Crane D, Garnett C, Michie S, et al. A smartphone app to reduce excessive alcohol consumption: Identifying the effectiveness of intervention components in a factorial randomised control trial. *Sci Rep* 2018; 8: 1–11.
32. Crane D, Garnett C, Brown J, et al. Factors influencing usability of a smartphone app to reduce excessive alcohol consumption: Think aloud and interview studies. *Front Public Heal* 2017; 5: 1–19.
33. Hamilton FL, Hornby J, Sheringham J, et al. DIAMOND (DIgital Alcohol Management ON Demand): A mixed-methods feasibility RCT and embedded process evaluation of a digital health intervention to reduce hazardous and harmful alcohol use. *Pilot Feasibility Stud* 2017; 34: 1–12.
34. Kitzinger J. The methodology of Focus Groups: the importance of interaction between research participants. *Sociology of Health & Illness* 1994; 16(1): 103–121.
35. Morgan DL. What do you get from focus groups? In: *The Focus Group Guidebook*. Thousand Oaks: SAGE Publications, Inc., 1998, pp. 55–64.
36. Krosnick J. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Appl Cogn Psychol* 1991; 5: 213–236.
37. Oppenheimer DM, Meyvis T and Davidenko N. Instructional manipulation checks: Detecting satisficing to increase statistical power. *J Exp Soc Psychol* 2009; 45: 867–872.
38. Babor TF, Higgins-Biddle JC, Saunders JB, et al. *The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care*, 2nd ed. Geneva: World Health Organisation, 2001.
39. Rollnick S, Heather N, Gold R, et al. Development of a short 'readiness to change' questionnaire for use in brief, opportunistic interventions among excessive drinkers. *Br J Addict* 1992; 87: 743–754.
40. West R. Time for a change: Putting the transtheoretical (stages of change) model to rest. *Addiction* 2005; 100(8): 1036–1039.
41. Kotz D, Brown J and West R. Predictive validity of the Motivation To Stop Scale (MTSS): A single-item measure of motivation to stop smoking. *Drug Alcohol Depend* 2013; 128: 15–19.

42. de Vocht F, Brown J, Beard E, et al. Temporal patterns of alcohol consumption and attempts to reduce alcohol intake in England. *BMC Public Health* 2016; 16: 917.
43. Qualtrics. Qualtrics Survey Software. http://www.qualtrics.com/ (2016).
44. Ponterotto JG. Qualitative research in counseling psychology: A primer on research paradigms and philosophy of science. *J Couns Psychol* 2005; 52: 126–136.
45. Braun V and Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006; 3: 77–101.
46. Garnett C, Crane D, West R, et al. Identification of behavior change techniques and engagement strategies to design a smartphone app to reduce alcohol consumption using a formal consensus method. *JMIR mHealth uHealth* 2015; 3: e73.
47. Petty RE and Cacioppo J. The elaboration likelihood model of persuasion. *Adv Exp Soc Psychol* 1986; 19: 123–205.
48. Oinas-Kukkonen H and Harjumaa M. Persuasive systems design: Key issues, process model, and system features. *Commun Assoc Inf Syst* 2009; 24: 486–501.
49. Hébert ET, Stevens EM, Frank SG, et al. An ecological momentary intervention for smoking cessation: The associations of just-in-time, tailored messages with lapse risk factors. *Addict Behav* 2017; 78: 30–35.
50. Collins RL, Morsheimer ET, Shiffman S, et al. Ecological momentary assessment in a behavioral drinking moderation training program. *Exp Clin Psychopharmacol* 1998; 6: 306–315.
51. Piasecki TM, Cooper ML, Wood PK, et al. Dispositional drinking motives: Associations with appraised alcohol effects and alcohol consumption in an ecological momentary assessment investigation. *Psychol Assess* 2014; 26: 363–369.
52. Yardley L, Morrison L, Bradbury K, et al. The person-based approach to intervention development: Application to digital health-related behavior change interventions. *J Med Internet Res* 2015; 17: e301.
53. Couper MP, Alexander GL, Zhang N, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. *J Med Internet Res* 2010; 12: e52.
54. Bossen D, Buskermolen M, Veenhof C, et al. Adherence to a web-based physical activity intervention for patients with knee and/or hip osteoarthritis: A mixed method study. *J Med Internet Res* 2013; 15: e223.
55. Deci EL and Ryan RM. Self-determination theory: A macrotheory of human motivation, development, and health. *Can Psychol* 2008; 49: 182–185.
56. Deci EL, Ryan RM and Koestner R. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol Bull* 1999; 125: 627–668.
57. O'Brien HL and Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. *J Am Soc Inf Sci Technol* 2008; 59: 938–955.
58. Brown E, Cairns P. A grounded investigation of game immersion. In: *CHI '04 Extended Abstracts on Human Factors in Computing Systems*. ACM, pp. 1297–1300.
59. Hassenzahl M and Tractinsky N. User experience: A research agenda. *Behav Inf Technol* 2006; 25: 91–97.
60. Hassenzahl M, Diefenbach S and Göritz A. Needs, affect, and interactive products: Facets of user experience. *Interact Comput* 2010; 22: 353–362.
61. Consolvo S, McDonald DW, Landay JA. Theory-driven design strategies for technologies that support behavior change in everyday life. In: *CHI '09 - Creative Thought and Self-Improvement*. 2009, pp. 405–414.
62. Milward J, Khadjesari Z, Fincham-Campbell S, et al. User preferences for content, features, and style for an app to reduce harmful drinking in young adults: Analysis of user feedback in app stores and focus group interviews. *JMIR mHealth uHealth* 2016; 4: e47.
63. Noar SM, Benac CN and Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol Bull* 2007; 133: 673–693.
64. Cooper A, Reimann R, Cronin D, et al. *About Face: The Essentials of Interaction Design*, 4th ed. Indianapolis, IN: John Wiley & Sons, Inc, 2014.
65. Yardley L, Bishop FL, Beyer N, et al. Older people's views of falls-prevention interventions in six European countries. *Gerontologist* 2006; 46: 650–660.
66. Wilsnack RW, Vogeltanz ND, Wilsnack SC, et al. Gender differences in alcohol consumption and adverse drinking consequences: Cross-cultural patterns. *Addiction* 2000; 95: 251–265.
67. Nolen-Hoeksema S and Hilt L. Possible contributors to the gender differences in alcohol use and problems. *J Gen Psychol* 2006; 133: 357–374.