Digital Behaviour Change Interventions to Break and Form Habits
Pinder, Charlie; Vermeulen, Jo; Cowan, Benjamin; Beale, Russell

DOI: 10.1145/3196830
License: None: All rights reserved

Document Version
Peer reviewed version

Citation for published version (Harvard):

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
© ACM, 2018. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in ACM Transactions on Computer-Human Interaction, (VOL25, ISS3, June 2018)
http://doi.acm.org/10.1145/3196830

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• Users may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 07. Aug. 2021
Digital Behaviour Change Interventions to Break and Form Habits

CHARLIE PINDER, University of Birmingham, UK
JO VERMEULEN, Aarhus University, Denmark
BENJAMIN R COWAN, University College Dublin, Ireland
RUSSELL BEALE, University of Birmingham, UK

Digital behaviour change interventions, particularly those using pervasive computing technology, hold great promise in supporting users to change their behaviour. However, most interventions fail to take habitual behaviour into account, limiting their potential impact. This failure is partly driven by a plethora of overlapping behaviour change theories and related strategies that do not consider the role of habits. We critically review the main theories and models used in the research to analyse their application to designing effective habitual behaviour change interventions. We highlight the potential for Dual Process Theory, modern habit theory and Goal Setting Theory, which together model how users form and break habits, to drive effective digital interventions. We synthesise these theories into an explanatory framework, the Habit Alteration Model, and use it to outline the state of the art. We identify the opportunities and challenges of habit-focused interventions.

CCS Concepts: • Human-centered computing → HCI theory, concepts and models; Ubiquitous and mobile computing design and evaluation methods; • General and reference → Surveys and overviews;

Additional Key Words and Phrases: Digital behaviour change interventions, behaviour change technology, persuasive technology, habit breaking technology, habit forming technology

ACM Reference Format:

1 INTRODUCTION

There is clear potential for digital behaviour change interventions (DBCIs), particularly those using pervasive technology, to change behaviour. Humans tend to make poor health decisions [Keeney 2008], while pervasive computing technology offering multiple detection and intervention points becomes cheaper and more widely owned.

The World Health Organisation (WHO) 2017 has identified four key lifestyle behaviours that impact severely on health: “tobacco use, unhealthy diet, lack of physical activity, and the harmful use of alcohol”. Meanwhile, technology becomes more ubiquitous. Smartphone ownership reached 80% in the UK in 2017 [Ipsos 2017], and there is rapid growth in wearables [Drake et al. 2017; Lee et al. 2015]. Smartphone health interventions are prevalent [Fiordelli et al. 2013; Klasnja and Pratt 2012; Lathia et al. 2013], and more ubicomp behaviour change interventions are emerging e.g. [Adams et al. 2015; Khot et al. 2015; Kim et al. 2016].

An obvious question is how best to design DBCIs to generate long-lasting results. The answer is not clear. This is primarily because the DBCI research area is cluttered with a large number of...
different theories, approaches and techniques, with research often failing to either evaluate long-term effectiveness or successfully identify long-term impact. Several meta-reviews have identified a general poor use of theory in DBCI research [Michie and Prestwich 2010; Wiafe and Nakata 2012], in DBCI apps [Cowan et al. 2012], and in behaviour change research in general [Michie and Johnston 2012; Sladek et al. 2006]. There are many theories and other explanatory structures, and they are generally not well applied. Crucially, several theories do not take the habitual nature of behaviours into account in their interventions, as we discuss in detail in Section 2.

We propose that a focus on changing habitual behaviours will improve the long-term efficacy of DBCIs. Our purpose is to outline the appropriate theoretical and strategic approaches to alter habitual and other automatic behaviours using technological interventions. We particularly focus on DBCIs using pervasive devices including smartphones, fitness trackers and smart home devices.

1.1 Article overview

In the first part of this article, Section 1, we establish why the study of habits is crucial for behaviour change and outline relevant domains for intervention. In Section 2, we give an overview of the main theoretical approaches in DBCI research to determine their suitability to target habits. We highlight three theories that together explain intervention points for habitual behaviours: Dual Process Theory, modern habit theory and Goal Setting Theory. In Section 3, we synthesise these theories into a framework (the Habit Alteration Model, HAM) to illustrate the main components and the potential intervention points for habit-focused DBCIs. We explore the related intervention strategies in Section 4, identifying the state of the art in the research field. We complement the strategies with a set of design principles in Section 5.

Section 6 outlines the main challenges in developing effective DBCIs to change habitual behaviour. In Section 7, we outline a future research agenda for the intervention opportunities afforded by increasingly pervasive context-aware technology. We present our conclusions in Section 8.

1.2 The importance of habits

1.2.1 What is a habit? Habits are learned impulses to perform a particular behaviour, triggered outside of conscious awareness by a particular context. Habitual behaviour is learned behaviour that is “frequently repeated, has acquired a high degree of automaticity, and is cued in stable contexts” [Orbell and Verplanken 2010]. Automaticity means habitual behaviours can be performed nonconsciously, i.e. “enacted with little conscious awareness” [Orbell and Verplanken 2010]. Note that in contrast to common usage of the word habit, we do not define it as the behaviour itself. Instead, following Gardner 2015, we define a habit as a link represented in associative memory between a certain context and a specific response. The occurrence of the context triggers a response impulse via context-response links. A habit is therefore a disposition to perform a given behaviour [Gardner 2015; Neal et al. 2006]. Habitual behaviour is the behaviour that results from this impulse.

Although habitual behaviours are triggered nonconsciously, people are not necessarily unaware of their actual behaviour. Instead, they tend to be unaware of the habit’s internal mechanisms [Stanovich 2005], such as the context-response associations [Wood and Rünger 2016]. This inability to introspect a habit’s underlying links makes them difficult to change: if the cause of an unwanted behaviour is not clear, then neither is the solution.

1.2.2 Habit prevalence and domains. Habits are highly prevalent and structure much of everyday life [Wood et al. 2014]. People report 43% of their behaviours as being performed without conscious thought ([Wood et al. 2002], study 2). Habitual behaviours span multiple domains: health [Gardner 2015], including eating [Robinson et al. 2013; Rothman et al. 2009; Wansink 2010], exercise behaviour [Aarts et al. 1997; Conroy et al. 2013] and physical activity [Rebar et al. 2016]; behaviour of
Digital Behaviour Change Interventions to Break and Form Habits

Table 1. Recent CHI habit mentions in full behaviour change papers.

<table>
<thead>
<tr>
<th>Year</th>
<th>Behaviour change</th>
<th>Mention habit</th>
<th>Substantive discussion</th>
<th>Behaviour domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHI ’15</td>
<td>36 (7%)</td>
<td>10 (2%)</td>
<td>4 [Hollis et al. 2015; Paay et al. 2015; Stawarz et al. 2015; Sugarman and Lank 2015]</td>
<td>Smoking; smartphone behaviour; unwanted behaviours; electricity consumption</td>
</tr>
<tr>
<td>CHI ’15</td>
<td>34 (6%)</td>
<td>17 (3%)</td>
<td>3 [Banovic et al. 2016; Hasselqvist et al. 2016; Sonne et al. 2016]</td>
<td>Childcare; sustainable transport; general routine behaviour</td>
</tr>
<tr>
<td>CHI ’17</td>
<td>37 (6%)</td>
<td>20 (3%)</td>
<td>2 [Banovic et al. 2016; Jansen et al. 2017]</td>
<td>General routine behaviour; eating behaviour</td>
</tr>
</tbody>
</table>

healthcare workers [Nilsen et al. 2012]; and environmental behaviours [Klöckner 2013]. Habits are also important in our use of technology [Bayer and Campbell 2012; Limayem et al. 2007], including participation in online communities [Wohn et al. 2012] and use of smartphones [Oulasvirta et al. 2012; van Deursen et al. 2015].

Despite a rise in interest in the habit construct in health psychology [Gardner 2015], few general behaviour change interventions currently use theory related to habit formation [Lally et al. 2008]. Likewise, few DBCI apps target habits [Stawarz et al. 2015]. Orji and Moffatt 2016 found that of 85 health domain DBCIs, only 3 targeted habits. Further, a review of 176 DBCI papers found only 11 that targeted nonconscious behaviour, and only 2 that mentioned related theory [Adams et al. 2015].

As a sample of recent DBCIs, we searched Google Scholar\(^1\) for behaviour change CHI papers from the past 3 years. The results are shown in Table 1. They demonstrate that few recent CHI DBCIs have targeted habits. For instance, out of 20 CHI ‘17 behaviour change papers that mentioned habit, only 2 address this directly as a construct in the research, with the vast majority only using the term colloquially.

DBCIs do, however, tend to focus on domains where habitual behaviours can emerge. Table 1 and Table 3 show that recent DBCIs have focused on health, technology and environmental behaviours alongside general routine behaviours. More broadly, Hamari et al.’s review of 95 DBCIs 2014 found that the main target domains were health (47%) and environmental behaviours (21%).

1.2.3 A challenge and an opportunity. Changing behaviour via habits represents both a challenge and an opportunity. The challenge is to break unwanted nonconscious habitual behaviours that are resistant to change. The opportunity is to use the habit mechanism to establish wanted nonconscious habitual behaviours that are similarly resistant to change. Habit formation can enable the maintenance of wanted behaviours [Sheeran et al. 2017b], since habitual behaviours are the default behaviour when people are unable or unwilling to make effortful decisions about how to behave [Neal et al. 2013]. They are performed automatically with little cognitive effort. These properties mean that DBCIs that can successfully form ‘good’ habits and break ‘bad’ habits are likely to have long-lasting behavioural effects [Lally et al. 2011; Rothman et al. 2009; Sheeran et al. 2013].

\(^{1}\)Search completed 13th June 2017. Search terms “habit”, (“behaviour change”“behavior change”and/or “persuasive technology”)
However, people often return to their unwanted behaviours over time [Bouton 2014]. This failure to sustain behaviour change is due to a lack of focus on automatic processes [Marteau et al. 2012] including habits. Behaviour change research tends to use deliberative interventions that rely on conscious resources: the provision of information is the most common DBCI technique [Webb et al. 2010b]. However, such interventions tend to be unsuccessful in the long term: Hillsdon et al. 2002 found evidence that over longer time periods, simply advising people to take more exercise is ineffective. Verplanken and Wood 2006 attribute such failures in part to the environment’s automatic activation of habitual behaviour to the detriment of people’s intentions. Reflecting this, habits are one of the key challenges for behavioural change policy [Jackson 2005]. There is therefore a clear research gap in considering habit as a key driver of long-term behaviour change. This article addresses this gap.

2 THEORIES OF BEHAVIOUR CHANGE

A critical component of addressing this gap is by understanding the key behavioural theories that contribute to our understanding of behaviour change at present and how they apply to habits. This section reviews the use of theory in behaviour change in general and in DBCIs.

2.1 Theory use in behaviour change research

The behaviour change research landscape is cluttered with many different theories, frameworks, models, techniques, strategies and patterns. Table 2 demonstrates the scale of the issue: just a few behaviour change researchers have identified tens of different behaviour change techniques, multiple ways behaviour might change, and numerous related theories and models.

Theory enables researchers to be more explicit about their assumptions, strategies and intervention targets [Rimer and Glanz 2005]. Despite—or perhaps because of—the number of competing models, there is a persistent lack of reference to theory in behaviour change research. The problem extends to DBCI research. Multiple reviews of DBCI research have found less than 50% specified a theoretical basis [Orji and Moffatt 2016; Wiafe and Nakata 2012]. Many interventions that claim to be based on theory fail to make explicit how the theory relates to the intervention or use theoretically predicted measures as evaluating criteria [Harris et al. 2011; Michie and Prestwich 2010]. Likewise, few interventions provide an in-depth understanding of the underlying mechanisms of behaviour and attitude change [Riley et al. 2011; Segerståhl et al. 2010], and few persuasive systems justify in detail their choice of behaviour change strategy, or the impact they are expected to have on their users [Foster et al. 2011].

This “theoretical gap” [Hekler et al. 2013] makes knowledge transfer between interventions difficult because it is not clear how and why a given intervention succeeds or fails [Nilsen 2015]. Under-use of theory is likely to impact the efficacy of the intervention, because important design characteristics are overlooked [Moller et al. 2017]. There is some evidence that interventions with a strong theoretical basis have a stronger association with efficacy [Taylor et al. 2012; Webb et al. 2010b], although this point is the subject of some debate [Michie and Prestwich 2010]. Ultimately, the paradigm shift to delivering behaviour change applications using technology is a key opportunity to deliver interventions based on systematic application of theory [Moller et al. 2017]. The gap in theory use reflects a lack of clarity around how to apply commonly-used theories to DBCIs. Health behaviour theories in particular have been criticised for being “woefully underspecified” [Sheeran et al. 2017b]. The inability of one single theory to address all aspects of behaviour change means researchers tend to use a ‘pick-and-mix’ approach to basing behaviour change strategies on theory [Bandura 1998; Hekler et al. 2013; Honka et al. 2011]. For example, the myBehavior system [Rabbi et al. 2015] incorporates elements from the Fogg Behaviour Model [Fogg 2009b], two decision theory models and Social Cognitive Theory, while Consolvo et al.’s set of design strategies 2009b
Table 2. Behaviour change patterns identified in the literature.

<table>
<thead>
<tr>
<th>Description</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of ways behaviour might change identified in the literature</td>
<td>35 [Fogg 2009a]</td>
</tr>
<tr>
<td></td>
<td>26 [Abraham and Michie 2008]</td>
</tr>
<tr>
<td>Number of behaviour change techniques, strategies or patterns</td>
<td>101 [Lockton et al. 2010]</td>
</tr>
<tr>
<td>identified in the literature</td>
<td>93 [Michie et al. 2013a]</td>
</tr>
<tr>
<td></td>
<td>56 [Wiafe and Nakata 2012]</td>
</tr>
<tr>
<td></td>
<td>50 [Lockton et al. 2009]</td>
</tr>
<tr>
<td></td>
<td>15 [Hamari et al. 2014]</td>
</tr>
<tr>
<td></td>
<td>7 [Fogg 2002]</td>
</tr>
<tr>
<td>Number of theories, frameworks or models identified in the</td>
<td>83 [Michie et al. 2014b]</td>
</tr>
<tr>
<td>literature</td>
<td>15 [Wiafe and Nakata 2012]</td>
</tr>
</tbody>
</table>

incorporates strands from Cognitive Dissonance Theory [Festinger 1957] and the Transtheoretical Model [Prochaska and Velicer 1997], amongst others.

In this pick-and-mix context, theory overlaps result in disagreement about which individual behaviour change strategy construct belongs to which theory [Doshi et al. 2003]. This is a particular problem for DBCI designers wishing to target habits because (a) it is not clear how the most commonly-used theories relate to habitual behaviour and (b) there is no theoretical consensus on habit mechanisms [Neal et al. 2006]. This article aims to provide clarity on the ability of the most commonly-used theories to explain habitual behaviours, and to bring together the most pertinent theories and strategies into an explanatory model that can inform the design of DBCIs to target habits.

2.2 Theory selection

We selected ten prominent theories in the literature and analysed their application to habit change: These theories either directly address habits (Behaviourism; Theory of Interpersonal Behaviour; Dual Process Theory; modern habit theory), are commonly used in behaviour change and DBCI research (Theory of Planned Behaviour; Social Cognitive Theory; Transtheoretical Model; Goal Setting Theory), directly addresses technology-mediated behaviour change (Fogg Behavior Model), or are very comprehensive in their coverage (COM-B). Table 3 shows a summary of the theories and an overview of their current use, the number of mentions in the ACM Digital Library within ACM Journals and Proceedings² (ACM search column), and citations for the key papers relating to each specific theory (Google Scholar and Scopus citations columns).

²Search terms: theory/model title and (“behaviour change” or “behavior change” or “persuasive technology”)
Table 3. Search results of theory/model mentions in the ACM plus citations and recent implementations

<table>
<thead>
<tr>
<th>Theory/Model</th>
<th>ACM</th>
<th>Google Scholar</th>
<th>Scopus</th>
<th>Citation source</th>
<th>Summary</th>
<th>Key determinants of behaviour</th>
<th>Recent Implementations</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theory of Planned Behaviour</td>
<td>56</td>
<td>50089</td>
<td>19954</td>
<td>[Ajzen 1991]</td>
<td>Behaviour is rational, determined by conscious intentions and Perceived Behavioural Control (an internal assessment of their ability to perform the behaviour).</td>
<td>Intention and Perceived Behaviour Control</td>
<td>[Bexheti et al. 2015; Comber and Thieme 2013; Schneider et al. 2016; Suh and Hsieh 2016]</td>
<td>Exercise; recycling; general behaviour change</td>
</tr>
<tr>
<td>Social Cognitive Theory</td>
<td>57</td>
<td>11673</td>
<td>2882</td>
<td>[Bandura 2001]</td>
<td>Behaviour is determined by an interaction between existing behaviours, the environment (including social factors), and personal cognitive, affective and biological influences.</td>
<td>Expected behavioural outcomes, environment and personal factors including self-efficacy.</td>
<td>[Khan et al. 2012; Rabbi et al. 2015]</td>
<td>Physical activity &amp; diet; snacking</td>
</tr>
<tr>
<td></td>
<td>264</td>
<td>N/A</td>
<td></td>
<td>[Bandura 1986]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

34 Operant conditioning results + 11 behaviourism results

4 search run 07/06/2017
### Table 3. Search results of theory/model mentions in the ACM plus citations and recent implementations

<table>
<thead>
<tr>
<th>Theory/Model</th>
<th>ACM</th>
<th>Google Scholar</th>
<th>Scopus</th>
<th>Citation source</th>
<th>Summary</th>
<th>Key determinants of behaviour</th>
<th>Recent Implementations</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trantheoretical Model</td>
<td>150</td>
<td>4942</td>
<td>2487</td>
<td>[Prochaska and Velicer 1997]</td>
<td>Six stages of behaviour, with ten processes for change. Interventions</td>
<td>Individual stage of change; self-efficacy;</td>
<td>[Paay et al. 2015; Park and</td>
<td>Smoking; smartphone use</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>can move people between stages by targeting self-efficacy and</td>
<td>decisional balance.</td>
<td>Gweon 2015; Southern et al.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>perception of advantages and</td>
<td></td>
<td>2017; Wittekind et al.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>disadvantages of behaviour.</td>
<td></td>
<td>2015</td>
<td></td>
</tr>
<tr>
<td>COM-B model</td>
<td>0</td>
<td>1123</td>
<td>612</td>
<td>[Michie et al. 2011]</td>
<td>An interacting system of capability, opportunity and motivation.</td>
<td>Capability, opportunity and motivation.</td>
<td>[Howlett et al. 2017; Lee et al.</td>
<td>Sleep; physical activity</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>948</td>
<td>214</td>
<td>[Fogg 2009b]</td>
<td>Behaviour is executed when 3 elements co-occur: motivation, ability and</td>
<td>Motivation, ability and a trigger</td>
<td>[Cambo et al. 2017; Lee et al.</td>
<td>Sleep; physical activity &amp; diet;</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>a trigger.</td>
<td></td>
<td>2017; Rabbi et al. 2015;</td>
<td>work breaks; energy conservation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sugarman and Lank 2015</td>
<td></td>
</tr>
</tbody>
</table>

5 search run 07/06/2017
6 plus 81 direct citations within ACM
Table 3. Search results of theory/model mentions in the ACM plus citations and recent implementations

<table>
<thead>
<tr>
<th>Theory/Model</th>
<th>ACM</th>
<th>Google Scholar</th>
<th>Scopus</th>
<th>Citation source</th>
<th>Summary</th>
<th>Key determinants of behaviour</th>
<th>Recent Implementations</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual Process Theory</td>
<td>6**</td>
<td>2960</td>
<td>1426</td>
<td>[Strack and Deutsch 2004]</td>
<td>Behaviour is determined by two distinct sets of cognitive processes: the Type 1 automatic system, formed of associative links; and the Type 2 conscious, deliberative system.</td>
<td>Interaction of two sets of cognitive processes: Type 1 processes (fast, automatic, nonconscious, associative); and Type 2 processes (slower, deliberative, conscious).</td>
<td>[Adams et al. 2015; Phelan et al. 2016; Pinder et al. 2015; Wang et al. 2014]</td>
<td>Eating; physical activity; privacy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13071</td>
<td>N/A</td>
<td>[Kahneman 2011]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal-setting theory</td>
<td>43</td>
<td>5258</td>
<td>2073</td>
<td>[Locke and Latham 2002]</td>
<td>Behaviour occurs where intentions are specified with an appropriate level of difficulty and specificity, and are accepted by users.</td>
<td>Intentions, contextual constraints</td>
<td>[Ding et al. 2016; Gouveia et al. 2015; Konrad et al. 2015; Lomas et al. 2017]</td>
<td>Stress; physical activity; learning</td>
</tr>
</tbody>
</table>

*Except where noted, ACM searches run 28/04/16. Google Scholar and Scopus searches run June 2017

*: theories that do not have one single article that defines them

7half of these papers were co-authored by the first author of the present paper
Table 3. Search results of theory/model mentions in the ACM plus citations and recent implementations

<table>
<thead>
<tr>
<th>Theory/Model</th>
<th>ACM</th>
<th>Google Scholar</th>
<th>Scopus</th>
<th>Citation source</th>
<th>Summary</th>
<th>Key determinants of behaviour</th>
<th>Recent Implementations</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A: items missing from Scopus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For theories with multiple sources (Behaviourism, Dual Process Theory, Social Cognitive Theory), we selected one or two relevant sources as a citation reference point. Since mere mentions and citations do not necessarily reflect implementations, we have augmented the results with recent applications of each theory from DBCI research (Recent Implementations column) and the domain of application (Domains column). Note that modern habit theory Section 2.7.2 is not listed in Table 3 because there is no one key paper that defines it and therefore it does not appear as a standalone search result.

This article does not aim to provide a comprehensive summary of all possible behaviour change theories and models that are available. Our aim is to consider the utility of applying the selected theories to changing habits and DBCIs, and to highlight recent research using them.

A key historical split in behaviour change theories and models is between behaviourism, which prioritise the role of the external environment in triggering behaviour, and cognitive theories, which argue that behaviour is also explained by abstract cognitive constructs such as thoughts and motivations. More recently, integrated theories and models have emerged to reconcile both standpoints since neither theory can account for all the complexities of behaviour change [Bandura 1998; Prochaska and Velicer 1997]. We consider our selected ten theories and models in each of these three categories below before addressing their overlaps and omissions.

### 2.3 Behaviourism

Behaviourism is a key theory in understanding habits because it focuses on the effects of the external environment on behaviour. It explicitly rejects the use of cognitive constructs to explain behaviour because they cannot be rigorously observed: only directly observable actions are considered.

#### 2.3.1 Overview

Behaviourists see habits as stimulus-response pairs formed outside conscious decision-making [West 2006] via two mechanisms of associative learning: classical and operant conditioning. A stimulus becomes associated with a particular response via repetition. Classical conditioning is the simple pairing of stimuli with responses; operant conditioning is the pairing of a stimulus-response with a positive or negative outcome. i.e. a reward for a wanted response, and a punishment for an unwanted one. Thus, rewarding a behaviour increases stimulus-response links and makes it more likely to be repeated. With repetition over time, any contextual cues that co-occur with a behaviour can trigger it [Davis 2001]. A behaviour is considered habitual when removing the reward does not diminish the behaviour, i.e. it is resistant to extinction. For example, a smoker who initially felt a positive reward from smoking (operant conditioning) may be prompted to smoke by the sight of a cigarette packet (classical conditioning), regardless of the subsequent reward. A key determinant of the impact of operant conditioning is how the reinforcement is delivered, or the reinforcement schedule [Staddon and Cerutti 2003]. A variable reinforcement schedule (where a reward is delivered to an average time or response rate, but not always at a given time or response) is the most effective in producing behaviour that is stable and resistant to extinction, rather than using constant or random reinforcements [Bijou 1957].

#### 2.3.2 Recent implementations and empirical evidence

Erev and Gopher 1999 argue that research suggests that a reinforcement learning model, where the probability of a certain behaviour being performed increases when it is positively reinforced, provides “an extremely good approximation of behavior in a wide set of situations” over alternative models that assume rationality. However, much of the research applying the principles of behaviourism to DBCIs is of a speculative nature. For example, Adams et al. 2009 suggest that pervasive exercise games are a good test-bed for empirical exploration of behaviourist learning principles but did not test this hypothesis.

Some researchers suggest that the use of operant conditioning and variable rewards in the technology domain underpin the use of social networks [Fogg 2009a], and the problematic use of
such social networks [Andreassen 2015] and the internet in general [Davis 2001]. However, neither claim is yet supported by empirical evidence.

There is some evidence that operant conditioning in the form of positive reinforcement may impact on unwanted habitual behaviours: a review of smoking cessation interventions during pregnancy found that strategies including the provision of incentives were the most effective [Lumley et al. 2009]. Positive reinforcements in the form of virtual rewards are common in DBCIs, but this is not a panacea for motivating behaviour, as we discuss further in Section 4.3.1. A recent implementation of variable reinforcement found some evidence that operant conditioning can change and maintain more secure behaviour, with a follow up period of 40 days, although the sample sizes were small [Villamarin-Salomón and Brustoloni 2010].

Negative reinforcement or punishment strategies are relatively rare in DBCIs [Kirman et al. 2010; Orji and Moffatt 2016]. This may be due to ethical concerns [Fogg 2002] and fear of disengaging users [Consolvo et al. 2009b]. Nevertheless, there is evidence that users are not necessarily deterred from interacting with DBCIs using aversive feedback [Foster et al. 2011]. One wearable DBCI that implements a punishment strategy is the Pavlok system, which allows users to trigger a mild electric shock (and/or beeps and vibrations) either manually or via sensors in order to punish a habitual behaviour they wish to break [Pavlok 2018], but the effectiveness of the device has yet to be demonstrated. Similarly, some researchers have implemented less painful punishment techniques, for example making interaction more tedious [Cowan et al. 2013; Cox et al. 2016; Foster et al. 2011], but none have been tested over the long term with a large user group.

2.3.3 Theoretical issues. A major theoretical criticism of behaviourism is its inability to explain higher-order behaviour involved in habits such as goals and conscious expectations of outcome. Behaviourism also cannot explain evidence that habitual behaviour can be triggered by cognitive constructs such as mood [Ji and Wood 2007]. Kihlstrom et al. 2007 argue that implicit learning involves some cognitive abstract representation of the knowledge, above and beyond the simple behaviourist associations.

2.4 Cognitive theory

Given the limits of behaviourism, we now turn to a key cognitive theory of behaviour change, the Theory of Planned Behaviour, which peers into the ‘black box’ of cognitive representations of external and internal behavioural drives.

2.4.1 Theory of Planned Behaviour.

Overview. The Theory of Planned Behaviour [Ajzen 1991] is a rational-action theory that specifies that intentions drive behaviour. A person’s behaviour is determined by their conscious intention to perform that behaviour and their Perceived Behavioural Control, an internal assessment of their ability to perform the behaviour. This intention is itself determined by behavioural attitudes, perception of subjective norms relating to the behaviour and Perceived Behavioural Control [Ajzen 1991]. It is “the most extensively studied social cognition theory” [Hardeman et al. 2002].

Recent implementations and empirical evidence. Schneider et al. 2016 applied the theory to unpick motivations of 643 mobile fitness coach users, finding that attitude, subjective norm and Perceived Behavioural Control were generally good predictors of intention, although the levels varied across personality types. The authors acknowledge the theory’s omission of possible nonconscious drivers of behaviour.

There is mixed evidence to support the theory from metareviews. Hardeman et al.’s review of 24 interventions 2002 found few studies actually using the theory to develop interventions, and a lack of evidence linking theory components to intervention outcomes. Webb & Sheeran’s
meta-analysis 2006 indicates that the intention construct is insufficient to fully explain behaviour change, with “a medium-to-large change in intention ... lead[ing] to a small-to-medium change in behaviour”. Crucially, an intention-behaviour gap persists, particularly in the presence of strong habits [Gardner et al. 2011; Webb et al. 2010b].

Theoretical issues. The Theory of Planned Behaviour is not a theory of behaviour change, and there is evidence that determinants of intention change over time [Suh and Hsieh 2016]. The theory thus has limited application to habits since they can only emerge in the presence of intentions that are enacted repeatedly in stable contexts. Sniehotta 2009 argues that the Theory of Planned Behaviour has major conceptual flaws, including no testable descriptions of how to modify intentions and therefore possibly behaviour. The theory omits context, habits and emotions as other possible determinants of behaviour [Jackson 2005; Sniehotta et al. 2014]. The inability of the model to deal with habits is particularly problematic: several studies [Triandis 1977; Webb and Sheeran 2006] have shown that interventions to alter intentions tended to impact on behaviour only where habits were not involved. Finally, assuming rationality where intentions are formed through conscious deliberative processes limits the theory’s ability to explain nonconscious behaviours.

2.5 Integrated models

Integrated models try to provide more overarching models of behaviour. They address dissatisfaction with the polarised view from behaviourists, that individual behaviour is solely determined by the environment, and cognitivists, that behaviour is solely determined by internal cognitive factors [Bandura 1978].

2.5.1 Theory of Interpersonal Behaviour.

Overview. The Theory of Interpersonal Behaviour [Triandis 1977] extends reasoned-action theories by explicitly including habitual behaviour and the context. Habit (expressed as frequency of past behaviour) and behavioural intention interact with situational conditions to determine behaviour. More frequently enacted behaviour weakens the intention-behaviour relationship. Behavioural intentions themselves are the product of attitudes, social factors and affect (the experience of emotion), with affect providing a largely nonconscious input into behavioural decision making. Thus the intention construct includes both nonconscious and conscious components.

Recent implementations and empirical evidence. There have been relatively few implementations of the theory. The few DBCIs based on it tend to focus on issues of technology acceptance. Moody & Siponen 2013 used the theory to explore the use of the Internet at work for non-work purposes, finding evidence to support the model’s key assumption that intention and habits both were significant in predicting the target behaviour. However, the research emphasised the social factors at play in the workplace, and it is not clear whether the results would generalise to other, less social domains. Gimpel et al. 2016 found that habit was a predictor of intention to use smartphones, although neither habits nor smartphone behaviour were measured.

Theoretical issues. Since the theory has been little used in DBCIs, it is difficult to establish its efficacy. This may be because it is not clear how to apply it to behaviour change interventions. Further, as we address in Section 6.2.3, using just frequency of past behaviour to approximate habits is insufficient. Danner et al. 2008 found evidence that past behaviour frequency only moderates the intention-behaviour relationship when information about context stability is also represented. Although the Theory of Interpersonal Behaviour includes “facilitating conditions”, it does not directly address the role of such conditions in forming habits – i.e. context stability. It may not therefore adequately capture habitual behaviour.
2.5.2 **Social Cognitive Theory.**

**Overview.** Social Cognitive Theory [Bandura 1978, 2011] states that behaviour is determined by an interaction between existing behaviours, the environment, and personal cognitive, affective and biological influences. Social influence is a particularly important environmental factor. Social Cognitive Theory suggests that behaviour change arises from two sorts of belief: firstly that a given response will have a desired outcome, and secondly that the individual believes themselves capable of the response [Clark and Janicic 2014]. The theory predicts that desired behaviours are performed where environmental barriers are low and self-efficacy is high [Armitage and Conner 2000]. It also incorporates elements of behaviourism via the mechanism of reinforcement for learned behaviours.

**Recent implementations and empirical evidence.** Implementations of Social Cognitive Theory tend to emphasise the key construct of self-efficacy rather than testing the theory as a whole e.g. [Rabbi et al. 2015]. Some empirical evidence supports the interaction between self-efficacy and behaviour change: a meta-analysis of physical activity studies found 3 techniques “associated with significant increases in both self-efficacy and physical activity behaviour; ‘action planning’, ‘reinforcing effort or progress towards behaviour’ and ‘provide instruction”’ [Olander et al. 2013]. Nevertheless, the overall effect was small, several other techniques had non-congruent effects on self-efficacy and physical activity, and the authors found that the reporting of intervention techniques was “inadequate” [Olander et al. 2013]. In addition, the identified techniques are consistent with Goal Setting Theory [Locke and Latham 2006], and behavioural theories of reinforcement, so it is not clear what additional contributions self-efficacy might provide, either at the theoretical or empirical level. Overall, Armitage & Conner 2000 argue that Social Cognitive Theory-based interventions typically account for small- to medium- levels of variance in behaviour.

**Theoretical issues.** Social Cognitive Theory has been criticised for failing to encompass habituation [Martin et al. 2014]. The theory relies on conscious, rational processing of behaviour change intentions and outcome expectancies, which do not reflect the observed automaticity of contexts triggering habitual behaviour. The implication is that habits are hard to change because they are perceived to be hard to change. The theory suggests that self-management is the key to breaking habits [Bandura 1998], but it is not clear how individuals can deal with low levels of deliberative cognitive resources to perform self-monitoring and self-regulation. Further, although the model does include context as a behavioural determinant, the focus is on the impact of social pressures such as role models and social support. Again, this omits the phenomenon of habits ceding control of behaviour to contextual cues. Social Cognitive Theory does integrate both cognitive and behaviourist elements, yet its hybrid nature makes it difficult to ascertain its efficacy and has led to calls for it to be subject to more rigorous testing [Martin et al. 2014].

2.5.3 **The Transtheoretical Model.**

**Overview.** The Transtheoretical Model [Prochaska and Velicer 1997] or “stages of change” health behaviour model was derived from a study of 872 people attempting to give up smoking on their own [Prochaska and DiClemente 1983]. The model identifies six stages of behaviour change, ranging from precontemplation (not even considering changing behaviour) to actively modifying their behaviour and/or the environment, through to maintaining the new behaviour and possible relapse. The model also identifies a set of ten processes for change, each of which has a different suggested emphasis for any given stage [Prochaska and Velicer 1997]. Prochaska & DiClemente 1983 argue that the key drivers of movement between the stages are self-efficacy (belief in one’s own ability to achieve a given goal) and decisional balance (weighing up pros and cons). This segmented model enables researchers to develop intervention strategies for each stage.
Recent implementations and empirical evidence. The Transtheoretical Model has been used widely in health behaviour change, e.g. see a meta-review of 71 empirical studies on the Transtheoretical Model and physical activity [Marshall and Biddle 2001], and in DBCIs in general, e.g. [Lin et al. 2006]. Many implementations focus on the precontemplation stage, where participants require information about their behaviour in order to motivate change e.g. [Park and Gweon 2015; Southern et al. 2017]. Wittekind et al. 2015 used the model in an anti-smoking DBCI to measure participants’ readiness to quit smoking rather than to design an intervention.

There are reasons to doubt the model’s efficacy. A metareviews of smoking cessation interventions – the domain that drove the model – found no evidence for a significant effect of interventions based on the Transtheoretical Model [Jepson et al. 2006]. Aveyard et al. 2009 concluded that there was “no evidence that Transtheoretical Model-based interventions [are] effective”.

Theoretical issues. The stages of change have been criticised as “arbitrary pseudo-stages” rather than genuine stages [Bandura 1998]. West 2005 lists several empirical challenges to the Transtheoretical Model, and argues that it should be discarded because it contains fundamental theoretical flaws. One key flaw in considering habits is that the Transtheoretical Model assumes that people make stable rational choices, rather than being subject to nonconscious influence such as impulses or habits.

2.5.4 COM-B model and the Behaviour Change Wheel.

Overview. The COM-B model [Michie et al. 2011] emerged from work in systematically reviewing and combining multiple behaviour change theories and frameworks relating to health behaviours. The model states that behaviour is determined by an interacting system with three essential components: capability, opportunity and motivation. Together with the Behaviour Change Wheel [Michie et al. 2014a, 2011], it provides comprehensive guidelines for behaviour change researchers to plan interventions. It essentially formalises the pick-and-mix approach.

The COM-B model addresses some gaps in rational-action models such as the Theory of Planned Behaviour by including nonconscious components like “impulsivity, habit, self-control, associative learning and emotional processing” [Michie et al. 2011]. The model includes both automatic and analytical processes in the motivation concept, which encompasses all brain processes that “energize and direct behaviour” [Michie et al. 2011], and is derived from the PRIME model [West 2006]. Opportunity includes all factors external to an individual that “make the behaviour possible or prompt it”, while capability includes all factors internal to an individual that contribute to their ability to perform a behaviour [Michie et al. 2011].

Recent implementations and empirical evidence. Walsh et al. 2016 used the COM-B model and Behaviour Change Wheel to devise an app for physical activity over 5 weeks. Their intervention group, which featured feedback and information about discrepancy between current behaviour and goal, had showed a small but significant improvement over the control. Lee et al. 2017 recently employed the model in a context-aware sleep intervention. However, COM-B was combined with many other techniques and strategies (e.g. Fogg’s Behaviour Model; goal setting theory; self-monitoring) and it is therefore difficult to make conclusions about the use of the model itself from their work.

Theoretical issues. One issue in using the COM-B model in DBCIs is that it is relatively new and therefore relatively untested in HCI [Cibrian et al. 2016]. The COM-B model and Behaviour Change Wheel are explicitly positioned as practical tools to design behaviour change interventions, rather than an explanatory theory, thus their specific application to habitual behaviour remains unclear. Further research is required to demonstrate how they can improve intervention efficacy [Michie
et al. 2011]. An important avenue of research is to explore how to use DBCIs to apply COM-B to dynamic contexts and individual preferences [Michie et al. 2013b].

Alongside more general psychological user models of behaviour, a number of models and methods have emerged that specifically relate user behaviour to technology. For reasons of brevity, here we outline Fogg’s highly-cited research into technology-mediated behaviour change.

2.5.5  Fogg Behavior Model.

Overview. Fogg’s work focuses on the concept of “captology”, or technology as a persuasive force in behaviour change [Fogg 2002]. This approach establishes seven key strategies that computers may use to change behaviour: reduction, tunnelling, tailoring, suggestion, self-monitoring, surveillance and conditioning; and three different roles the computer may play in these strategies: as a tool, as media and as a social actor. The Fogg Behavior Model [Fogg 2009b] grew from this focus on technology and practical strategies. It is a general cross-domain model that proposes behaviour has three key determinants: motivation, ability and a trigger. All these elements must occur at the same time to generate a particular behaviour.

Fogg also created the Behavior Grid [Fogg 2009a], a taxonomy of 35 ways behaviour might change; and the Behavior Wizard [Fogg and Hreha 2010], which attempts to merge the previous two items; and the “Tiny Habits” model ([Fogg 2015] not yet published as a peer-reviewed model but already cited e.g. [Daskalova et al. 2017; Kuo and Horn 2017], involving changing the performance environment, breaking the required habitual behaviour into small steps and rewarding small step completion.

Recent implementations and empirical evidence. Establishing empirical evidence for the Fogg Behaviour Model is difficult because it is a process model rather than a theory. Researchers also tend to use it alongside other models, e.g. Cambo et al.’s work 2017 uses the Fogg Behavior Model with the Health Action Process Approach, which alongside a small sample size and short intervention period (1 day) makes it difficult to draw conclusions about the model. Sugarman & Lank 2015 used the model to inform a qualitative survey of methods to encourage reduction in electricity consumption, although again they incorporated elements from other theories e.g. operant conditioning.

Theoretical issues. Fogg’s work is essentially a behaviour change principles approach [Noar et al. 2008]. The Fogg Behaviour Model is an attractively simple conceptualisation of behaviour with clear design implications: provide people with an appropriate trigger, motivation and ability to perform a wanted behaviour and it will occur. However, the underlying psychological mechanisms of change are less clear. Further, although Fogg sees the point of persuasive technology as “fundamentally about learning to automate behavior change” [Fogg 2009b], and his Behavior Wizard [Fogg and Hreha 2010] attempts to simplify behavioural repetition to form new habits, the Fogg Behavior Model has little to say about the automatic components of behaviour or routine behaviour [Oulasvirta et al. 2012]. Ferebee 2010 found ambiguities and a lack of guidelines in mapping existing interventions to the Behavior Grid.

2.6  Discussion of competing models

2.6.1  Theory Overlaps. Health behaviour change models contain many components, some of which are shared or overlap [Taylor et al. 2006]. For example, intentions are a key behavioural determinant in the Theory of Planned Behaviour, the Theory of Interpersonal Behaviour and Social Cognitive Theory, although the theories differ in the elements that determine that intention. The notion of Perceived Behavioural Control from the Theory of Planned Behaviour is similar to parts of the Theory of Interpersonal Behaviour’s notion of facilitating conditions and Social Cognitive Theory’s concept of self-efficacy. Common to these models is the implicit assumption that users form
intentions along rational, conscious lines. Most health behaviour theories assume that conscious attitudes and intentions, self-efficacy and social influences impact most on behaviour [Noar et al. 2008].

The Transtheoretical Model, the Fogg Behaviour Model and COM-B are process models, which focus on translating research into concrete behaviour change interventions rather than more theory oriented framework models [Nilsen 2015]. Framework models, by contrast, focus on explaining the determinants of specific outcomes [Nilsen 2015]. Process models may have greater utility to intervention designers than framework models, since the latter may omit specific implementation detail [Rogers 2004]. However, process models may not always highlight the underlying theory: although the COM-B model is explicitly couched in the PRIME model of motivation [West 2006], the Fogg models do not have explicit theoretical underpinnings.

There are also clear overlaps between the three-part models that address interactions between behaviour, the environment and internal cognitive factors (Fogg Behavior Model, COM-B, Theory of Interpersonal Behaviour and Social Cognitive Theory). Indeed, Lee et al. 2017 recently combined the Fogg Behavior Model and COM-B into a single approach which assumed a given behaviour occurred when “opportunity, ability, motivation, and a trigger all align”.

One key common determinant of behaviour is intention (and related motivations), which we explore here in a little more detail.

**Intentions & motivations.** An intention is a decision to undertake a particular behaviour at a future point in time. An intention encompasses the person’s motivation to perform the behaviour - the direction (to perform the behaviour or not) and the intensity (how much value they assign to that performance) [Sheeran 2002]. The Theory of Planned Behaviour, the Theory of Interpersonal Behaviour and Social Cognitive Theory all assume that intentions are key determinants of behaviour, and those intentions arise from the likelihood and desirability of the outcomes of a given behaviour [Deutsch and Strack 2010; Webb and Sheeran 2006]. However, it is not clear exactly how intentions drive behaviour [Bruin et al. 2012], nor how the theories see any interaction between habit and intentions.

Motivation is a value attached to a particular intention. Motivation is a key construct in the Theory of Interpersonal Behaviour, Social Cognitive Theory, the COM-B model and the Fogg Behavior Model, and is central in moving people from contemplation to active stages in the Transtheoretical Model. We agree that DBCIs require people to be consciously motivated to change their behaviour to engage in interventions at the very least at the outset. However, most models (with the key exception of COM-B) focus on conscious, rational motivations, omitting important automatic aspects of motivation including the impact of contextual cues, internal physiological states (e.g. hunger) and emotions. These automatic elements are specified in the PRIME theory of motivation [West 2006], on which COM-B is based.

**Theory gaps.** Despite some consensus on behavioural determinants between the theories, there are gaps in their ability to drive the design of behaviour change interventions. Not all the models are dynamic, or specify how their constructs or the relationships between them change over time. This limits their application to DBCIs that can adapt rapidly to their users and to changing inputs [Riley et al. 2011]. Some models omit the impact of the context, habit and/or emotions in determining behaviour.

There is little consensus on how to combine the overlapping constructs to change behaviour [Noar et al. 2008]. In particular, not all the theories explore how behaviour changes. For example, the Theory of Planned Behaviour predicts behaviour, rather than addressing how intentions can change over time [Suh and Hsieh 2016]. The exception is the Transtheoretical model, which explores when intentions change, but not how [Armitage and Conner 2000]. Many theories incorporate elements
of behaviourist operant conditioning, e.g. the use of positive reinforcement in the Transtheoretical Model and Social Cognitive Theory [Adams et al. 2017]. However, despite the crucial role the environment plays in behaviourism, many common theories emphasise individual/interpersonal variables rather than broader social/environmental variables [Davis et al. 2015; Taylor et al. 2006]. This is a crucial omission in their application to understanding and changing habits. Since contextual features are crucial to both triggering and forming habits, models that omit them are unlikely to capture habitual behaviour determinants adequately.

Habit itself is a key construct omitted from many behaviour theories, particularly health-related theories [Nilsen et al. 2012], despite compelling empirical support for its role as a moderator of the intention-behaviour link [Sheeran et al. 2017b; Webb and Sheeran 2006]. In general, models that assume a rational, deliberative process as a key determinant of behaviour (e.g. the Transtheoretical Model and the Theory of Planned Behaviour), are insufficient to explain habitual behaviour given the persistence of an intention-behaviour gap in the presence of strong habits [Gardner et al. 2011; Webb et al. 2010b].

Theories that include habit either mention it somewhat in passing (COM-B model) or restrict its determinants too narrowly (e.g. behaviourism’s failure to incorporate cognitive constructs that operate during habit formation). Even theories that explicitly incorporate habit e.g. the Theory of Interpersonal Behaviour and Social Cognitive Theory, fail to explain how and why habits operate, which limits their practical application. Further, although COM-B, the Theory of Interpersonal Behaviour and Social Cognitive theory all include elements of nonconscious motivation, it is not clear how the conscious and nonconscious elements work together to determine behaviour.

A good candidate for filling this theoretical gap in DBCIs is Dual Process Theory. Despite Dual Process Theory being the “probably one of the most significant theoretical developments in the history of social psychology” [Gawronski and Creighton 2006], it has been little used in DBCIs [Adams et al. 2015; Orji and Moffatt 2016; Webb et al. 2010a]. The under-use of Dual Process Theory is significant because together with modern habit and goal theory, it directly addresses the intention-behaviour gap and allows us to address the research gap in understanding how to build habit-focused DBCIs. One reason for this under-use is because of a lack of a clear, practical framework to apply Dual Process Theory to DBCIs. Providing this is the motivation for the current paper.

2.7 Bridging the theory gaps

In the next section, we outline three theories that address habit formation and habit breaking: Dual Process Theory, modern habit theory and Goal Setting Theory. We argue that bringing these three theories together can bridge the research gap in understanding how to change habits using technology and therefore move to close the intention-behaviour gap. Dual Process Theory allows us to see how conscious and nonconscious forces interact to determine behaviour; modern habit theory indicates how these might combine to determine habitual behaviour; and Goal Setting Theory informs effective goal-setting strategies to help drive habit formation through behavioural repetition.

2.7.1 Dual Process Theory.

Overview: Dual Process Theory argues that behaviour emerges from two distinct sets of processes: Type 1 (broadly automatic, e.g. habits) and Type 2 (broadly conscious, e.g. behavioural intentions). Type 1 processes are nonconscious cue-driven, heuristic, impulsive, associative, contextual, automatic, parallel processes that operate at speed; while Type 2 processes are conscious goal-directed, slower, rational, considered, rule-based, abstract serial processes [Evans and Frankish 2009; Evans 2011]. This split roughly maps onto the behaviourist/cognitivist rationalist divide, with habits
forming part of the Type 1 set. Not all the nonconscious, automatic behaviours triggered by Type 1 processes are habitual [Marteau et al. 2012]. People may act in line with an impulse in response to a cue or in line with nonconscious goals without the action becoming a stable, repeated behaviour (see Section 4.1.2 on Priming).

The crucial difference between behaviourism and Dual Process Theory is that habits rest on cognitive constructs, and thus may be altered using both cognitive and behavioural techniques. Dual Process Theory thus unites the behaviourist-cognitivist divide: behaviour is the outcome of an interplay between both Type 1 and Type 2 processes.

We have outlined the common assumptions of Dual Process Theory, but there is no one definitive version of Dual Process Theory [Evans 2008]. It is a family of theories emerging from multiple fields of research [Stanovich 2011], including Borland’s CEOS model 2013; 2016, the Reflective-Impulsive model from social psychology [Strack and Deutsch 2014], and System 1-System 2 theory from behavioural economics [Kahneman 2011].

Recent implementations & empirical evidence. Few DBCIs currently use Dual Process Theories [Adams et al. 2015]. Orji and Moffat’s review of 85 DBCIs since 2000 found zero implementations [Orji and Moffatt 2016]. However, DBCI implementation pilots are emerging. Examples include our pilot on smartphones [Pinder et al. 2017] and two studies by Adams et al. 2015 in “mindless computing”. Phelan et al. 2016 recently used the theory to inform a qualitative investigation into privacy behaviour, finding that the dual process view helps to inform the “privacy paradox” where users’ privacy behaviour is not consistent with their privacy concerns.

Behaviour change research as a whole has recently begun to advocate the targeting of Type 1 processes alongside Type 2 approaches [Bargh and Morsella 2010; Dolan et al. 2012; Marteau et al. 2012; Sheeran et al. 2013]. Dual Process Theories are being used increasingly in health behaviour interventions [Hofmann et al. 2008], e.g. Kremers et al. 2006 used Dual Process Theory to derive a practical framework to explore the impact of environmental factors on weight gain.

In terms of neuroscientific evidence to support the theory, action-outcome behaviour (cognitivist goal-directed behaviour) and context-response behaviour (behaviourist habits) are associated with two different sets of brain processes [Gasbarri et al. 2014; Graybiel 2008; Yin and Knowlton 2006]. Habits as context-response links are therefore somewhat distinct from their origins as action-outcome sequences [Maia 2009]. At an experimental level, Pressseau et al. 2014 measured deliberative and automatic predictors for six different healthcare behaviours (e.g. providing weight advice; prescribing for diabetes) using questionnaires and found that both types predicted behaviour.

Theoretical issues. Since Dual Process theories are still little-used in DBCIs, it is difficult to establish their efficacy. Further, implementing Dual Process Theory is not trivial: there are multiple versions, and researchers are still actively developing the theory as it applies to behaviour change [Borland 2016; Wiers et al. 2013]. Nevertheless, regardless of the specific Dual Process Theory chosen, they agree on two key predictions: behaviour is an outcome of both Type 1 and Type 2 processes; and Type 1 processes (including habits) will dominate when Type 2 resources are depleted, during distraction, high cognitive load, time pressure, adverse mood and low self-control [Hofmann et al. 2008; Muraven and Baumeister 2000]. The relative importance of Type 1 and Type 2 processes as a determinant of behaviour also varies with personality [Sladek et al. 2006]. Thus the influence of Type 1 and Type 2 processes on an individual’s behaviour will vary both over time and in comparison with other people.

Dual Process Theory helps us to understand the underlying mechanisms of translating cues into action. However, it does not in itself provide a practical framework of applying the theory to habit-targeting DBCIs. For this we need to examine two additional theories: modern habit theory and Goal Setting Theory.
2.7.2 Modern habit theory.

**Overview.** Modern habit theory also integrates both stimulus-response behaviourist theories and goal-directed cognitive reasoned-action theories, e.g. [Wit and Dickinson 2009]. We have outlined the key points and some empirical evidence for the existence of habit in Section 1.2. This section therefore addresses a recent implementation and some theoretical issues.

**Recent implementations.** Stawarz et al. 2015 investigated the formation of habits in-the-wild, focusing on the relatively simply behaviour of reporting what participants had for lunch. In a 4-week study, they found that automaticity (as measured by the Self-Report Behavioural Automaticity Index, outlined in Section 6.2.3) was hindered both by smartphone reminders and positive reinforcement [Stawarz et al. 2015].

**Theoretical issues.** Habit research is ongoing across multiple fields, with ongoing challenges in determining the exact mechanisms underlying habit formation [Tobias 2009; Yin and Knowlton 2006] and in studying its automaticity [Gasbarri et al. 2014]. Nilsen et al. 2012 suggest that there is a lack of empirical evidence supporting interventions based on habit formation theory [Stawarz et al. 2015].

2.7.3 Goal Setting Theory.

**Overview.** We include Goal Setting Theory [Locke and Latham 1990, 2002, 2006] because it explicitly explores how best to form goals to drive behavioural repetition when Type 2 processes predominate to support habit formation. It fills the theoretical gap in many theories lacking detail in how to specify intentions. The theory proposes that goals must be accepted by users to be effective, that feedback on goal progress is important, and that two key aspects of goal setting determines their efficacy: difficulty and specificity. Hard, specific goals are more effective than easy, vague ones. Contextual constraints are considered to be a moderator [Latham et al. 2017]. The original Goal Setting Theory focused on conscious goals, but research is moving towards incorporating nonconscious dimensions [Latham et al. 2017]. This move is partly in response to ongoing debate about the theoretical underpinnings of nonconscious goal priming, which we discuss in Section 4.1.2.

**Recent implementations.** DBCI researchers often employ Goal Setting Theory alongside other behaviour models and theories to augment their interventions. For example, Ding et al. 2016 used goal setting theory predictions within the design of a context-aware walking app based on the Fogg Behavior Model. They found some qualitative evidence that users liked short-term step goals rather than daily or weekly goals, but it is unclear whether the results can be generalised.

**Empirical evidence.** Meta-analysis indicates that specific, difficult goals improved performance compared to asking people to “do their best”, with effect sizes from .42 to .80 [Locke and Latham 1990], cited in [Locke and Latham 2002]. However, there is evidence that who assigns the goal makes a difference: when the DBCI sets the goals, easier goals may be more effective. Lomas et al. 2017 examined learning game goals and found that when goals were self-selected, moderately difficult tasks were most motivating, whilst when they were externally assigned, easiest games were most motivating. Konrad et al. 2015 found evidence that adaptive, easy goals set by user’s technology were more motivating than difficult goals in a month-long experiment with 65 participants.

**Theoretical issues.** The goal setting theory picture is complicated by a lack of consensus on the measurement of goal commitment [Hollenbeck et al. 1989] or indeed goal difficulty. When designing DBCIs with Goal Setting Theory, researchers should also take individual characteristics into account. Orji et al. 2017 found evidence that goals in line with goal setting theory are motivators for people high in extraversion and conscientiousness.
2.8 Summary
We have outlined a series of theoretical approaches to changing habits. Three theories in particular are good contenders to fill the theoretical gap in explaining habitual behaviour: Dual Process Theory, modern habit theory and Goal Setting Theory. To ease interpretation of how these three theories combine to address habits, we have brought them together in a conceptual framework to explore habit intervention points for DBCIs.

3 THE HABIT ALTERATION MODEL

3.1 Overview
The Habit Alteration Model (HAM) is a practical conceptual model that synthesises Dual Process Theory, modern habit theory and Goal Setting Theory so it can be applied more easily to DBCIs. A model’s function is to provide a descriptive simplification of a phenomenon [Nilson 2015]. The HAM provides a conceptual, theory-driven graphical simplification of how these theories suggest that external and internal factors combine to generate both habitual and non-habitual behaviour. Its purpose is to provide a practical tool to describe and assess habit-targeting DBCIs, as called for by several researchers [Aarts et al. 1997; Hollands et al. 2016]. It allows researchers to devise new interventions that do not solely rely on limited deliberative cognitive resources to change behaviour.

The HAM is shown in Figure 1. Behaviour is a function of: (1) the context consisting of a set of cues; (2) Type 1 associative processes relating cues to behavioural impulses; (3) Type 2 deliberative processes generating explicit intentions; and (4) individual differences (e.g. impulsivity), which determine the relative impact of Type 1 and Type 2 processes on behaviour. The model is dynamic: at the Filter stage, cues flow through perception and Type 1 and Type 2 attention filters to create an input set. At the Prepare stage, Type 1 and Type 2 memory processes match these cues to potential responses, Type 1 impulses or Type 2 intentions. These compete to become a single response at the Act stage. Information from observed response and outcomes feed back into the model and therefore into both Type 1 and Type 2 processes. Solid lines indicate processes that run continuously; dashed lines indicate processes that may run. Note that the HAM is not intended to represent the various highly complex physical architectures that operate in the brain. Instead, we present it as a virtual conceptual architecture where the boundaries between the systems need not be rigid [Sloman 2002].

With sufficient repetition of simple behaviours in stable contexts, cycles around the Filter-Prepare-Act stages become more automatic, and the corresponding context-response habitual behaviour links become stronger and proceed faster. People’s behaviour then transfers from slower Type 2 to faster Type 1 processes, from the conscious right-hand side of Figure 1 to the nonconscious left-hand side. Habit disruption strategies aim to call on Type 2 deliberation to override automatic Type 1 processes. Although disruption can be employed to a user’s advantage, e.g. in error checking or reducing technology over-use [Cox et al. 2016], disruption also makes behavioural outcomes less stable because its success depends in part on available cognitive resources.

Dual Process Theory predicts that any behaviour may be the result of the simultaneous influence of Type 1 and Type 2 processes [Kremers et al. 2006; Presseau et al. 2014]. Impulses to respond in a habitual way triggered by a given context compete with other impulses, and with intentions from Type 2 processes, to determine a response [Gardner 2015]. The dominant response is determined by the relative strength of the items on the Potential Response stack, and is influenced by cognitive resources and an individual’s cognitive capacity and processing style [Sladek et al. 2006].
Next we outline in more detail how the HAM models how habits are triggered, then how they are formed.

### 3.2 Habit-trigger process

#### 3.2.1 Filter. We start with a set of cues that make up a given context. These cues include external features of the environment such as physical locations and other people, and internal features such as mood or physiological drives such as hunger [Wood et al. 2014]. This broad set of all possible cues in a given context is first ‘filtered’ by perception processes. They are then again filtered via Type 1 implicit attention processes (F2) and, optionally, Type 2 explicit attention processes (F3). Type 1 implicit attention filtration is influenced by mood, attitudes and stereotypes such that some cues
receive preferential implicit attention than others [Deutsch and Strack 2010]. Type 2 processes in the deliberative system may also use directed attention (F3) to select specific cues from the context. This top-down conscious attention has limited cognitive resources [Norman and Shallice 1986], so its ability to select cues is impacted by cognitive load. The end result of the filter process is a subset of cues as inputs to the potential response generation process, Prepare.

3.2.2 Prepare. The cue inputs are used by both Type 1 and Type 2 memory processes to generate behavioural schemas for action [Strack and Deutsch 2014]. Type 1 processes (P1) generate impulses from implicit memory while Type 2 processes (P3) generate intentions from explicit memory. These separate schemas compete to become enacted behaviour, via a mechanism to integrate competing, parallel inputs into a single behaviour [Bargh and Morsella 2010]. We represent this mechanism in the HAM as a Potential Response stack (P6).

Type 1 memory processes are fast, modular and parallel, so multiple impulses may be generated by the set of available cues and placed on the Potential Response stack. Not all automatic impulses that arise from Type 1 are habits. Habits are context-response impulses for behaviour that has been repeated in a stable context. Automatic goal impulses are instead goal-response links, allowing for nonconscious goal-driven behaviour when a goal acts as a cue [Aarts and Dijksterhuis 2000]. Impulses emerging from Type 1 memory processes may also be of a simple approach or avoid type [Keatley et al. 2013], for example instinctive behaviour to flinch from a loud sound. We explore the application of both goal-related and instinctive impulses in DBCIs in Section 4.1.2.

The Potential Response stack (P6) may also contain conscious, deliberative intentions arising from Type 2 deliberative memory processes (P3). These may arise from explicit goals via the mechanisms of self-control (P4) and self-monitoring (P5). Intentions that have been set using Goal Setting Theory are assumed to take priority on the stack. Intentions include the intention not to act, i.e. impulse stifling as part of self-control (P4), if an unwanted impulse is likely to be enacted. This ability requires that the contents of the Potential Response stack are to a certain extent accessible to conscious thought [Bargh and Morsella 2010]. Self-monitoring (P5) may also place intentions on the stack in line with goals in response to processing information from self-tracking.

The order of impulses and intentions on the Potential Response stack is determined by several factors: match with the particular cue [Norman and Shallice 1986]; affect towards the cue and/or response [West 2006]; and accessibility [Danner et al. 2008; Kahneman 2003]. Placing value on degree of ‘match’ means that behaviour enacted more often appears higher on the stack than less previously-enacted behaviour, since the match with particular cue will be stronger. Thus an impulse to perform a behaviour that has been repeated in a stable context will appear higher in the Potential Response stack: these are habits.

3.2.3 Act. A competitive winner-takes-all process determines which single behaviour is performed from the competing schemas on the Potential Response stack [Hofmann et al. 2009]. Any potential response (impulse or intention) that exceeds a certain act threshold (the pink dashed line in P6 in Figure 2) will be enacted if there are no rival potential responses [Wood et al. 2014]. Where competing potential responses cross the act threshold, arbitration using Type 2 processes is required, as detailed below.

Following a response, there may be a particular outcome, for example a corresponding change in the environment or a reward. Information on the response and outcome feed back into implicit and explicit memory processes and therefore may impact on subsequent Act phases [Sun et al. 2005; Wood and Neal 2007].

Act arbitration process. Figure 2 shows a zoomed-in view of different possible states of the Potential Response stack with respect to the act threshold (the pink dotted line), and the resulting
Fig. 2. Response outcomes from different Potential Response stack states: single impulses and intentions above the act threshold will be enacted, otherwise arbitration will occur. Responses may occur regardless of the presence of intentions - in state 1 and state 4, intentions may not be present (indicated by unfilled intention item). Height order indicates relative value in the stack.

- **State 1**: Where a single impulse crosses the act threshold, the impulse will be enacted regardless of intention: intention D may or may not be present.
- **State 2**: Where a single intention is strongly-held such that it alone crosses the act threshold, its target behaviour will be enacted regardless of competing, weaker impulses.
- **State 3**: Where conflicts between a Type 1 impulses and a Type 2 intention occur with similar implicit values so that both cross the act threshold.

behavioural response. Impulse A, impulse B and impulse C are impulses to perform behaviours A, B and C respectively; intention D is an intention to perform behaviour D. The relative value of items on the Potential Response stack is indicated by height order: the higher in the stack, the higher the relative value.
threshold, Type 2 processes may be alerted to arbitrate [Wood and Neal 2007]. State 4 shows that arbitration may also be alerted to differentiate between competing impulses, regardless of intention.

Arbitration is the implicit core of many habit-breaking strategies. These strategies try to populate the Potential Response stack with conscious Type 2 intentions to compete with unwanted other potential responses in order to trigger deliberative arbitration. However, calling on Type 2 arbitration resources imposes cognitive load, which is not always available. Where arbitration cannot be performed, the most likely response is the highest-value impulse in the stack. Arbitration is hampered by multiple load factors including other Type 2 processes, ego depletion and time pressure and individual factors including working memory capacity and low trait self-control [Hofmann et al. 2009]. This explains why many effortful intentions to change behaviour fail: when Type 2 cognitive resources are low, default Type 1 impulses predominate [Hofmann et al. 2009].

In case 3, if cognitively costly arbitration cannot be carried out, impulse A will predominate because impulses appear more quickly in the Potential Response stack in excess of the act threshold than slower intentions. Where impulse A represents any Type 1 habit, generated by repeating a simple behaviour in response to stable contextual cues, habits will predominate when cognitive resources are low.

3.3 Habit formation process

How might people form habits such that the default behaviour is their desired behaviour? Repetition is key. Habit formation requires that a given response (Act stage) is repeated in a stable context, i.e. with a stable set of cues arising from the Filter stage. With repetition, the impulse to perform the given response emerges as highest in the Potential Response stack (Prepare stage), triggered by the stable cues. The response (Act stage) may then proceed without conscious attention, i.e. the intervention of Type 2 processes. In behaviourist terms, stimulus-response links have been established.

This habit formation process can be accelerated by rewarding the required response, i.e. operant conditioning, providing a rewarding outcome (A2). Rewards can promote the learning of context-response links [Wood et al. 2014]. A reward does not have to be explicit for a habit to form. For example, Conroy et al. 2013 found evidence for an habitual element in sedentary behaviour despite this behaviour not being explicitly rewarded or even consciously intended. The difficulties of providing rewards in the right way to form habits are addressed in Section 4.3.1.

The key task for DBCIs that seek to change habits is to foster behavioural repetition in a stable context. The traditional behaviour change intervention for habit forming is to use conscious Type 2 processes (the right-hand side of Figure 1) to drive repetition via mechanisms of self-monitoring, reminders and self-control. However, as we outline below, the HAM also allows this repetition to be targeted via nonconscious Type 1 means.

3.3.1 Role of emotion. Most behaviour change theory does not include emotion as a determinant of behaviour [McDonald 2014]. Emotions primarily impact on behaviour through indirect mechanisms [Baumeister et al. 2007], although emotion and mood impact on both Type 1 and Type 2 attention and memory processes [Bargh and Morsella 2010; Deutsch and Strack 2010].

At the Filter stage, emotions may act directly as cues in both Type 1 and Type 2 attention processes. Emotional state, or mood, may also shape how these attention processes work by biasing attention to both negative or positive cues [Fox et al. 2009]. At the Prepare stage, emotion –either conscious or nonconscious– can impact the order of items in the Potential Response stack [West 2006] and thus influence behaviour. Negative emotions and stress can also limit deliberative cognitive processing ability and therefore limit Type 2 interventions [Tice et al. 2001; Wood et al. 2014].
Finally, in the Act stage, affect can also serve as a reward, for example how ‘enjoyable’ or ‘exciting’ an exercise behaviour is [Kaushal et al. 2017].

3.3.2 Example. To illustrate the HAM model, consider the unwanted behaviour of snacking whilst watching television. In the Filter stage, the context (F1) includes the physical environment (e.g. sofa, television), the presence of the snacks themselves, and the snacker’s current mood. While the snacker’s conscious Type 2 attention (F3) is focused on the television programme, the snacker’s nonconscious Type 1 attention (F2) is aware of the presence of the snacks via sight, feel and smell. Therefore cues of snack presence and mood form inputs to the Prepare stage. During this stage, these cues feed into Type 1 memory processes (P1) and generate an impulse to eat the snack, which appears on the Potential Response stack (P6). Assuming there are no competing impulses or intentions, and given sufficient previous repetition of the snacking behaviour in the presence of these cues, then the snacking impulse will exceed the act threshold. Since Type 2 cognitive resources are occupied in watching the television programme, levels of self-monitoring (P5) and self-control (P6) to stop snacking are low; the person does not place a counter-snack intention on the Potential Response stack (P6). The response at the Act phase is snacking behaviour. The outcome (A2) may be some intrinsic reward from the snack (implicit feedback).

4 STRATEGIES FOR DBCIS TO INTERVENE IN HABITUAL BEHAVIOUR

This section discusses the strategies our framework identifies to intervene in habitual behaviour based on habit literature. The HAM is used to illustrate the intervention targets. The analysis covers habit breaking and habit making strategies; some techniques are appropriate for both. For habit-forming strategies, the key question is how to move from Type 2 to Type 1 processes, from right to left in Figure 1. This is a movement from behaviour arising from slow, limited, serial, explicit systems to behaviour arising from faster, parallel implicit systems. For habit-breaking strategies, the key question is how to alter existing Type 1 processes without relying on cognitively effortful disruptive Type 2 resources. Points of intervention for habit-changing DBCIs are denoted by numbers F1-F3, P1-P6 and A1-A2 in Figure 1. The following section sets out a variety of different strategies to target these points of intervention.

4.1 Filter (Phase 1)

Removing or avoiding a specific cue that forms part of a cue-response link in an unwanted habit will mean that the undesired response is not initiated or therefore performed. This approach is challenging because the associative links in Type 1 processes are not available to introspection [Neal et al. 2012; Orbell and Verplanken 2015]. An individual is therefore unlikely to be aware of which cues are relevant to their unwanted habitual behaviours. An alternative strategy is to introduce cues that are likely to trigger required responses.

4.1.1 Alter context (F1).

Outline. The rationale for altering the context is to add or remove cues in order to affect which impulses and intentions arise in the Potential Response stack. Alterations may include changing cue properties such as ambience and size and/or placement, e.g. proximity and availability [Hollands et al. 2013]. Context alteration is suggested as particularly applicable in the unhealthy eating domain [Wansink and Chandon 2014]. With our unhealthy snacking example, a DBCI could suggest to the user to not buy the snacks in the first place, or suggest replacing them with a healthy snack whilst watching television.

Evidence. The primary implementations of context-altering DBCIs are ambient persuasive technology systems. They are designed to change behaviour and/or attitude unobtrusively by exerting
influence on people without requiring their focal attention [Ham et al. 2009]. Examples include altering a workspace to encourage people towards the stairs [Rogers et al. 2010] and augmenting a shopping trolley to influence consumer behaviour [Kalnikaite et al. 2011]. However, there is a lack of a strong evidence base for the efficacy of specific context-altering strategies [Hollands et al. 2013].

An additional context alteration strategy is the alteration of moods. There is some evidence to suggest that small mood-altering interventions can be successfully ported from psychology labs onto smartphones to alter moods [Meinlschmidt et al. 2016], although the technique has yet to be applied to DBCIs to alter habits.

**Challenges.** Determining which cues to alter is not trivial. Kremers et al. 2012 identified 35 broad environmental changes to promote change in food and activity behaviours, and the individual efficacy of such changes remains unclear. It is a particular challenge to detecting appropriate emotional cues for a given behaviour. We address the challenges of context detection further in Section 6.1. Large-scale ambient persuasive interventions can have high installation costs. This cost drawback has triggered research into altering “micro-environments”, contexts on a smaller scale, for example product labelling or design [Hollands et al. 2013].

4.1.2 **Priming (F1, P1, P6).** An alternative context-alerting strategy is delivering cues via technology that users carry as part of their personal context, e.g. smartphones. Type 1 processes include associative memory links between cues and affective and behavioural responses [Strack and Deutsch 2014]. They can be activated using priming, the unobtrusive presentation of cues to activate relevant mental representations [Shalev and Bargh 2011]. Priming can increase accessibility of a goal concept [Kahneman 2003], making it more likely to be performed [Bargh et al. 2001]. Positive valence towards a concept can also increase its accessibility [Kahneman 2003; West 2006], providing opportunities for affective priming [Custers and Aarts 2007]. In the HAM, priming is providing a specific cue within the context (F1) that crosses the attention barrier to form an input to Type 1 Prepare processes (P1). These processes select a target impulse from memory. Ideally, this impulse is relatively high on the Potential Response stack (P6) and therefore likely to be enacted. If enacted and repeated in a stable context, then it will become a habitual behaviour [Wood and Neal 2007].

DBCIs may implement two forms of priming behaviour: the activation of instinctive paths to achieve certain behaviour, or the activation of learned constructs such as goals.

**Instinctive paths.**

**Outline** Several ‘instinctive’ context-response paths already exist within humans that may be used to prime behaviour change. Stanovich 2005 argues that these fast “genetic goals” are more easily primed than learned goals, and have the advantage of being more universal. Evidence of these instinctive paths include the influence of auditory [Spence and Shankar 2010] or other environmental cues [Wansink 2010] on eating behaviour, and the influence of apparent monitoring (displaying images of eyes) increasing compliance with honesty boxes [Bateson et al. 2006]. The latter example is particularly interesting because it implies the possibility that merely the act of appearing to monitor may be sufficient to increase compliance to some behavioural norm. With our unhealthy snacking example, a DBCI could display a prime indicating monitoring (e.g. an image of an observer), to encourage more positive snacking behaviour. This example also highlights issues of ethics, which we address further in Section 5.7 and Section 6.3.
Several pervasive systems have implemented instinctive triggers targeting difference senses, particularly in the fitness domain. Several sound-based DBCIs react to user heart rate by selecting [Nirjon et al. 2012], altering [Oliver and Flores-Mangas 2006] or auto-generating workout music [Bauer and Waldner 2013], while the Zombies, Run! [zombiesrungame 2018] app exploits the 'flight from fear' hard-wired instinct by cueing users to run faster using sounds of raving hordes of zombies. An alternative sound-based system alters the user’s walking sounds in order to change their perceived body weight and therefore change their gait [Tajadura-Jiménez et al. 2015].

One sight-based intervention, in the healthy eating domain, is the Mindless Plate. This prototype explored whether perceptions of food portion size could be altered using coloured plates, with somewhat encouraging short-term results [Adams et al. 2015]. Smell-based interventions are starting to emerge based on evidence that this approach can change behaviour over a week even when administered during a single night’s sleep [Arzi et al. 2014]. Amores & Maes 2017 developed the Essence prototype necklace that uses smell as a nonconscious influence on mood and cognitive performance. An alternative ‘instinctive’ path implemented in DBCIs is that of social priming, e.g. [Aharony et al. 2011]: the idea that humans are predisposed to react to the cue of seeing another person perform a behaviour by responding with similar behaviour. This theory is supported to a certain extent by research in neuroscience [Kessler et al. 2006], but evidence for efficacy is mixed [Froehlich et al. 2010]. Nevertheless, new technology affords the opportunity to test large-scale social contagion in real-world settings through social networks, with some evidence that exercise is socially contagious across such structures [Aral and Nicolaides 2017].

Instinctive primes are a good candidate for research where deliberative strategies have repeatedly failed, for example in the healthy eating domain [Obrist et al. 2014; Pels et al. 2014; Wansink 2010]. Indeed, using primes within a food-consumption context to achieve changes in eating may be easier than changing people’s generally poor ability to monitor their consumption [Wansink and Chandon 2014].

The key challenge in using instinctive primes is to identify the correct prime for a given behaviour. Once identified, the prime needs to be delivered in a sufficiently salient manner such that it crosses the implicit and/or explicit attention filters (F2, F3). If the primed behaviour is not repeated in a stable context, no habit will be formed. This is not just an issue of context detection, since enacting the desired behaviour is not guaranteed, given differences in individual responses and concurrent different states of the Potential Response stack (P6). Further research is required to determine how best to deliver the instinctive prime such that the related impulse appears at the top of the Potential Response stack P6, and is therefore likely to be enacted in the absence of competing intentions or arbitration (see Section 3.2.3). The technique is likely to be most successful to direct people during situations of high cognitive load (e.g. driving, working).

If we can prime instinctive associations, to what extent can we prime learned associations, such as goal constructs, i.e. an association between a goal and the behaviour required to achieve the goal [Danner et al. 2011], to drive habit formation? This strategy may mitigate some of the challenges in identifying and delivering instinctive cues outlined above. Modern goal research indicates that goals, instead of definitively forming part of conscious deliberation in Type 2 processes, can not only be activated nonconsciously [Aarts et al. 2008; Stajkovic et al. 2006], but also operate nonconsciously [Chartrand and Bargh 1996; Förster et al. 2007; Pessiglione et al. 2007]. Priming goals results in more persistent accessibility of the related concepts than simply priming behaviour alone, at least until the goal-related behaviour is enacted [Bargh and Morsella 2010]. Primed impulses will therefore
have more value in the Potential Response stack (P6). With our unwanted snacking example, a DBCI could support the user by displaying the prime of a photo of themselves consuming an alternative, wanted snack, or even more abstract representations of goals of being thinner such as Giacommetti sculptures [Stämpfli and Brunner 2016].

Evidence  Chen et al. 2014 found some evidence in a single session pilot that priming intentions increased user engagement in an exergame, while Custers & Arts 2007 showed that goal priming increased both valence and affective valence and impacted on effort to pursue a goal. The priming can also be delivered nonconsciously, e.g. subliminally. Caraban et al. 2017 applied subliminal priming in a browser plug-in by decreasing the opacity of key words and found some evidence of priming on subsequent item selection. Our own research in the area found mixed results of the immediate impact of subliminal priming across different primes (polygons, words, photos) delivered on smartphones [Pinder et al. 2017]. Evidence for the efficacy of goal priming as a behaviour change technique in general is mixed. Some evidence supports the approach [Sheeran et al. 2013, 2017b], while other evidence is more ambivalent [Wood and Neal 2007].

Challenges  Goal priming requires work with users ahead of the intervention to instil the goal along Goal Setting Theory lines such that it can be primed. Priming with no pre-training implicitly assumes that participants already associate the target cue with the goal. Target cues need to be designed carefully to avoid ironic effects where instructions such as ‘do not X’ primes behaviour X [Earp et al. 2013]. There are some theoretical objections: Papes 2016, in line with the COM-B model, argues that goal priming procedures can only be successful where individuals also hold sufficient motivation, capability and knowledge to pursue it. We agree that successful priming depends on the appropriate selection of primes and construction of related goals [Ruijten et al. 2011] such that the goals cross the conscious /nonconscious divide in memory, P2 in Figure 1. Nonconscious goal priming experiments in psychology frequently use supraliminal tasks where the aim of the task is concealed – e.g. by tasking users with a word search where the target words relate to the goal of ‘performing well’ [Bargh et al. 2001]. However, word search or scrambled sentence tasks are difficult interventions for DBCIs using pervasive computing technology, particularly using small-screened or unobtrusive technology. Subliminal priming provides a possible alternative method of prompting nonconsciously. Subliminal priming can help to avoid user irritation [Ham and Midden 2010], and may be more likely to result in goal-related behaviour than conscious prompts [Glaser and Kihlstrom 2005]. We suggest further research into the use of subliminal goal priming to increase the accessibility of related behaviour, and thus increase the likelihood of the goal-related behaviour being enacted [Custers and Aarts 2007; Pinder et al. 2017].

Priming activation is distinct from habit activation. A habit is a learned context-response impulse, while priming activates multiple mental concepts in memory related to the prime [Wood et al. 2014]. To prime habit formation, DBCIs either need to be sufficiently context-aware to prime only within a stable context (to form new habits) or accept that their intervention may need to be persistent if stable-context-priming is not possible. The latter technological dependency may not lead to long-term behaviour change if the technology is abandoned [Renfree et al. 2016].

Further research is required to determine the most effective method of priming. Priming design choices include design of the prompt, duration, repetitions required [Pinder et al. 2017] and delivery mechanism, e.g. using opportunistic interventions tied to an unlock screen on a smartphone such as in [Pinder et al. 2016b] or during browsing [Caraban et al. 2017].

4.1.3 Alter cue salience (F2, F3).
Outline. The likelihood that a given cue gets through the implicit perception filter is determined by its salience. Thus, a more relapse-resistant strategy is to reduce the salience of contextual cues for unwanted responses, whilst also increasing the salience of cues for wanted responses, using Cognitive Bias Modification (CBM) techniques for attention biases [MacLeod et al. 2009], CBM-A. An attention bias is the tendency for a given cue to receive disproportionate implicit and/or explicit attention, points F2 and F3 respectively in Figure 1. Reducing this attention for unwanted cues then reduces the resulting unwanted response because the cue is less likely to become an input to Type 1 and Type 2 memory processes, and likewise the reverse with wanted cues. With our unhealthy snacking example, a DBCI could try to reduce attention bias for the snack by giving the user a serious game to pair images of their problematic snack with something they find revolting.

Evidence. Biases affecting attention can be altered by appropriate training [Hertel and Mathews 2011] and evidence from psychology labs of the potential for the technique to change behaviour is encouraging, e.g. in the healthy eating domain [Kakoschke et al. 2014]. However, there are relatively few DBCI implementations, and the ones that exist are primarily at pilot stages [Pinder et al. 2016a,b]. One randomised controlled trial porting CBM-Attention techniques onto smartphones found inconsistent results with only small effects on attention bias scores [Enock et al. 2014].

Challenges. CBM techniques face challenges in identifying the relevant cues that need increasing/decreasing in salience (see Section 6.1) and ensuring longevity of the newly-learned responses [Hertel and Mathews 2011]. The mixed empirical evidence indicates that additional research is required to determine how best to port these techniques from psychology labs to DBCIs.

4.2 Prepare: target the context-response link (Phase 2)

4.2.1 Train context-response (P1, P6).

Outline. The main technique to train context-response links is Cognitive Bias Modification for approach biases, CBM-Ap. An approach bias exists when an individual has a default action or impulse towards an unwanted cue, i.e. a Type 1 Prepare processes, P1 in Figure 1. For example, a smoker may have an approach impulse towards a cigarette. CBM techniques targeting approach biases train individuals to inhibit responses or reject these unwanted items, and to accept alternative wanted items. For example, the smoker might be trained to reject cigarettes and accept chewing gum, or with our unhealthy snack example, be trained to reject the unhealthy snack and accept a fruit alternative.

Evidence. Two CBM-Ap studies have found small but significant results following brief training with challenging participants and a long follow-up period. Wiers et al. 2011 trained alcoholics with 4x15 minutes lab training sessions. Participants used a joystick to push away images of alcoholic drinks on a desktop display, and pulled towards them images of non-alcoholic drinks. The training was sufficient to alter the intervention group’s small approach bias for alcohol to a strong avoidance bias, reflected in marginally significant differences in relapse rates between intervention and control groups [Wiers et al. 2011]. Wittekind et al. 2015 trained psychiatric inpatients with a similar anti-smoking CBM-Ap over 4 sessions and found small but significant differences in self-reported nicotine consumption at 3-month follow-up.

Challenges. As with CBM-A, it is not yet clear how best to translate CBM-Ap from controlled conditions in the lab to pervasive technology in the wild. Cue identification may also be an issue, although it is easier for people to identify their unwanted approach biases than unwanted attention biases. The behavioural impacts of the two CBM-Ap studies was small, and intervention groups also received standardised Type 2 interventions (e.g. motivational interviewing). Nevertheless, evidence
of impact of the minimal Type 1 training indicates potential in using more pervasive technology to
deliver larger numbers of training sessions in situ [Pinder et al. 2016b]. CBM-Ap has parallels with
automating self-control, which focuses on response inhibition, as discussed in Section 4.2.5.

4.2.2 Implementation Intentions (F2, F3, P1, P3).

Outline. This approach specifically tries to bridge the gap between explicit Type 2 intentions
and implicit Type 1 impulses. Implementation intentions are specific, concrete if-then plans that
link particular if contexts (i.e. sets of cues) to a desired response, then. They aim to automate the
then behaviour by delegating its control to the selected contextual if [Gollwitzer et al. 2005].
Implementation intentions are therefore special form of automated goals that can bridge the
intention-behaviour gap [Webb and Sheeran 2006; Wood and Rünger 2016], and are argued to be
a good strategy for habit formation apps [Stawarz et al. 2015]. Implementation intentions are
conscious Type 2 processes that emulate Type 1 context-response habits. With our unhealthy snack
example, a DBCI could support the user to form and rehearse an intention to counter the unwanted
snacking, for example ‘if I am watching TV, then I will only snack on apple slices’.

The mechanism through which implementation intentions work is increased accessibility [Webb
and Sheeran 2008], such that the resulting behavioural intention to perform the then response is
highest in the Potential Response stack (P6). Through rehearsal, sufficiently concrete and relevant
implementation intentions can become impulses, moving from Type 2 deliberative processes into
Type 1 automatic processes [Einstein and McDaniel 2005]. Implementation intentions fit with the
persuasive strategy of rehearsal found in several behaviour change models e.g. [Fogg 2002], but
have a stronger theoretical grounding that fits with Dual Process Theory and modern habit theory.

Evidence. The evidence for implementation intentions is good: a meta-analysis found that they
“had a positive effect of medium-to-large magnitude (d = .65) on goal attainment” [Gollwitzer
and Sheeran 2006]. However, the meta-analysis did not consider whether the behavioural goals
related to habitual behaviour or not, and other literature reviews note a heterogeneity in effect sizes
[Hagger and Luszczynska 2014]. Prestwich et al. 2010 examined the use of both goal-based and
plan-based SMS to boost implementation intentions to undertake daily brisk walking, finding that
both conditions significantly increased the target behaviour compared to a control group, although
the study was based on self-report.

DBCIs can support implementation intentions by providing support for intention formation and
rehearsal. DBCIs can enable even inexperienced users to quickly learn if-then plans with multiple
triggers or actions [Ur et al. 2014]. In-situ reminders [Tobias 2009] and rehearsal [Veling et al. 2014]
can support the accessibility of implementation intentions.

Challenges. An effective DBCI using implementation intentions would ideally be able to advise
users on which contextual cues are appropriately stable cues (i.e. ifs) on which to base their if-
then plans, together with appropriate reminders and rehearsal. However, difficulties of accurately
monitoring context cues and behaviour, which we detail in Section 6.1, mean implementation is
challenging. The evidence for habit breaking using implementation intentions is more mixed:
some research suggests they are not good at controlling strong habits [Wood and Rünger 2016],
while Sheeran et al. 2017b suggest that they have been successful in the smoking domain. Further
research is still required into the effectiveness of technology-mediated implementation intention
formation and rehearsal.

4.2.3 Provide information (P3).

Outline. The provision of information is common in both DBCIs [Pejovic and Musolesi 2014;
Webb et al. 2010a] and in behaviour change interventions in general, e.g. in healthcare [Nilsen et al.
Here we define the provision of information as providing the user with data intended to alter their conscious decisional balance, e.g. to give them statistics on positive health outcomes for a given desired behaviour. The underlying “information gap hypothesis” with this approach [Cowan et al. 2013] implicitly assumes a rational choice model (e.g. the Theory of Planned Behaviour), where people will alter their conscious behavioural intentions to counter a given behaviour in the light of the information provided. With our unhealthy snacking example, a DBCI could provide the user with data on the adverse health effects of regularly eating unhealthy snacks.

Evidence. Providing information can in some circumstances change behaviour, albeit with a small impact: a meta-analysis of public information campaigns showed a weighted mean effect size of .05 on behaviour [Anker et al. 2016]. However, in applying information to the DBCI field, the evidence is mixed. Comber & Thieme 2013 used information provision as a strategy to counter habitual recycling behaviours with a just-in-time recycling-monitoring system. They found no impact of their awareness-raising on attitudes or on behaviour in qualitative feedback of 22 participants over 5 weeks. It is also difficult to evaluate its use in DBCIs because it is rare to find interventions that use the provision of information as a single tactic.

Crucially, the long-term effects of providing information on behavioural intentions are not stable. A randomised control trial (RCT) with longitudinal research (12 months) demonstrated that a strategy of advising people to do more exercise is ineffective in the long term [Hillsdon et al. 2002]. Likewise, within the eating domain, there is evidence of a ‘mindless eating’ gap of 15-20% of consumption, which persists within stable environments regardless of an individual’s knowledge of nutritional information [Bellisle et al. 2004; Wansink and Chandon 2014].

Challenges. There are three key problems with providing information as a behaviour change strategy. Firstly, deliberative cognitive resources may not be available such that Type 2 processes can attend to, analyse and deliberate the information. Interventions often use disruption alongside information provision to demand Type 2 resources to process the information they provide [Verplanken and Wood 2006]. Secondly, people may not change their attitudes and/or behaviours in line with the information. Combor & Thieme 2013 found that although disruption (an audio signal) alerted their users to the need to attend to their behaviour, they were unsuccessful at providing the right information to change the behaviour in the desired way. Thirdly, Type 1 processes may bias the information itself as an input to Type 2 processes. For example framing effects, such as presenting the same information in either positive or negative ways, impact subsequent Type 2 judgements [Kahneman and Tversky 2000]. Nevertheless, this presence of cognitive bias in decision making can also present an opportunity. Lee et al. 2011 argue that varying the way in which information is delivered –specifically in providing the required behaviour as the default –can be successful, and this is a key tenet of choice architecture or ‘nudge theory’ [Thaler and Sunstein 2008]. DBCI designers therefore need to be aware that the way in which information is presented may also have an impact, beyond simple Type 2 decisional balance effects. Indeed, reframing a suggested option as a default, rather than outlining its advantages, is closer to a priming intervention than simple provision of information.

4.2.4 Just-in-Time reminders (F3, P3).

Outline. An alternative to simply advising people to change their behaviour is to leverage pervasive context-aware technology to provide just-in-time reminders to behave in a particular way [Moller et al. 2017]. To distinguish such reminders from alternatives (e.g. response priming or implementation intentions), we define just-in-time reminders here as direct, specific behavioural suggestions delivered at the expected point of enactment. Just-in-time reminders are applicable to habit formation (advising people to repeat a wanted behaviour in a stable context) and habit breaking
(advising people to refrain from an unwanted behaviour in a given context). With the unhealthy snacking example above, a user’s phone could alert them to the unwanted eating behaviour and suggest an alternative.

**Evidence.** Although reminders can support the development of habits, they can have a diminishing effect [Tobias 2009]. They can also prompt reactance, where users try to reassert behavioural control from a perceived threat to it, (discussed in detail in Section 6.4.2), particularly where users are instructed to suppress thoughts of an unwanted behaviour [Palfai et al. 1997]. As with priming, if users are not paying full attention to the reminders, ironic effects may result [Earp et al. 2013]. For example, a DBCI that warns ‘do not eat your favourite snack’ may inadvertently trigger the user to eat that snack. Further, even if the ‘correct’ response is performed, this may be dependent on the presence of the technology as a cue as part of the context-response link itself. This makes the new habit more fragile and susceptible to disruption [Renfree et al. 2016]. Stawarz et al. 2015 found evidence in a 4-week trial that electronic reminders increased behavioural repetition but impeded automaticity. Without automaticity, once the DBCI is removed, the behaviour is unlikely to persist.

Researchers have also made efforts to reduce the complexity of reminders to reduce their cognitive load. Ding et al. 2016 tried to identify low-disruption incidental points to deliver walking prompts using context-aware technology. Nevertheless, they found that users reported embarrassment at inappropriate suggestions.

**Challenges.** Behavioural repetition in a stable context is crucial to forming habits. Therefore just-in-time habit-forming reminders must be context-aware. However, there are persistent technological issues with context detection (see Section 6.1), and few approaches to capturing behaviour also capture causal relationships between context and response [Banovic et al. 2016]. This limits their ability to remind in a habit-forming way.

From a theoretical perspective, Wood & Neal 2007 suggest that attracting attention to a given response may inhibit automated habit formation in favour of automatic goal pursuit. This in turn may affect the long-term sustainability of the required response; if a DBCI withdraws its reward for goal-directed behaviour, the response will cease. By contrast, a behaviour that is successfully habituated will persist as long as the set of trigger cues does not contain the intervention technology itself.

### 4.2.5 Train self-control (P4).

**Outline.** Self-control is the ability to “alter [your] own behavioral patterns so as to prevent or inhibit [the] dominant response” [Muraven and Baumeister 2000]. From a Dual Process Theory perspective, it is considered to be part of Type 2 processes [Metcalfe and Mischel 1999], although there is evidence that it can, through repetition, become automated into Type 1 processes where ‘stop’ behaviours are consistently mapped onto the same cue(s) [Verbruggen and Logan 2009]. Self-control could therefore provide a mechanism for people to resist habitual impulses to act in unwanted ways. Taylor et al. 2013 suggest that computer-based training to enhance self-control could take advantage of neuroplasticity to play a role in treatment for drug addiction alongside pharmacological treatment. Webb12 et al. 2010b express surprise that few interventions against addictive behaviour have used self-control strategies. With our unhealthy snacking example, a DBCI could be designed to support the user to train themselves to resist the snacks by using go/no go serious games [Lawrence et al. 2015].

**Evidence.** De Ridder et al. 2012 found evidence for a relatively large relationship between self-control traits and habits and suggest that self-control is important in both making and breaking habits. Adriaanse et al. 2014 by contrast suggest that self-control operates mostly through breaking
unwanted habits. Verbruggen & Logan 2009 used 5 studies to show that participants can be trained to automatically inhibit responses to unwanted items by practicing inhibition. Muraven 2010 argues that self-control training is generalizable, such that the mechanism can be improved by training small acts of conscious behavioural inhibition, regardless of either of domain or whether the subjects believed it would help. This strategy may therefore provide an important lever for behaviour change for circumstances when Type 2 processes are in control.

Some self-control DBCIs are starting to emerge. Cranwell et al. 2014 found that a 3-daily, 4-week training task significantly increased self-control test scores in an intervention group compared to a control group. Two sets of research have found some impact of an internet-delivered go/no-go food image task on weight loss [Lawrence et al. 2015; Veling et al. 2014]. Lawrence et al. 2015 trained participants to inhibit responses to unwanted foods with 4 sessions over 1 week and found a significant, medium sized drop in self-reported calorie intake. The study also reported that 88% of participants would continue the training if shown to be effective. The theory is that practising inhibition of impulses to eat via the training can impact on eating decisions, particularly for those who are overweight.

Nevertheless, it is not clear precisely how the self-control mechanism works, nor is the evidence for its efficacy consistent: Miles et al. 2016 found no effects on either Type 1 or Type 2 self-control following a 6-week self-control training programme.

Challenges. The key challenge is that effortful (Type 2) self-control is unlikely to be effective in the long term to change habitual behaviour because of a limit on deliberative cognitive resources [Wood and Neal 2007]. Where Type 2 self-control fails, old habits will re-emerge. The inability to introspect habits [Hagger et al. 2015] may also hamper attempts to limit unwanted behaviour through self-control where people are unaware of their habit cues. A strong association between affective state and self-control capacity [Economides et al. 2015; Tice et al. 2001] may also hamper behavioural persistence. However, repeated practice of self-control can itself become automatic and therefore part of Type 1 processes, similar to CBM-Ap training [Fishbach and Shah 2006]. As with automating other behaviour, repeating the same self-control or impulse-resisting control in stable contexts should ease the transition of self-control into automatic processes, moving from Type 2 to Type 1 memory processes. The new ‘response’ of self-inhibition must be repeated sufficiently such that this becomes the default response rather than the old, unwanted behaviour.

Effortful self-control has limited capacity [Baumeister 2002] and may even have ironic effects [Wegner 1994] prior to becoming automatic. Therefore, as with other Type 2 processes, it is unlikely to operate under conditions of high cognitive load, and may even have the opposite effect to that intended.

4.3 Act: target the habit response (Phase 3)
Targeting the response aspect of the context-response link may involve: the use of self-monitoring to reveal previously unknown response patterns; or operant conditioning on the response outcome (A2).

4.3.1 Revalue outcome (A2).

Outline. The key strategy to revalue outcomes is the providing rewards for ‘correct’ behaviour [Gouveia et al. 2015] or punishments for ‘incorrect’ behaviour [Kirman et al. 2010]. From a behaviourist standpoint, this is operant conditioning, the most common application of which is contingency management—to provide a positive outcome for a given behaviour to strengthen the context-response link [Maia 2009]. Rewards are not a necessary part of habit formation, for example
when the target behaviour is intrinsically rewarding [Lally et al. 2008], but they may accelerate its development into automaticity.

From a cognitivist standpoint, rewards increase the expected value of a given behaviour. It is the subject of some debate whether rewards operate on a nonconscious Type 1 behaviourist level or a conscious Type 2 cognitivist level [Capa et al. 2011; Neighbors et al. 2013]. Using the HAM model in Figure 1 to illustrate, rewards boost the position of the matching *impulse* (nonconscious Type 1) or *intention* (conscious type 2) on the *Potential Response* stack (P6). Virtual rewards are commonly used in DBCIs [Hamari et al. 2014; Orji and Moffatt 2016], although punishment strategies are much rarer [Kirman et al. 2010], as we outlined in Section 2.3.2. Such rewards are mostly designed to attract attention, which means they may be overlooked in conditions of high cognitive load. With our unhealthy snacking example, the user could reward themselves for consuming healthy snacks as an alternative, for example by transferring a small amount of money into a virtual jar for each healthy snack consumed, and/or punish themselves for consuming unhealthy snacks by giving a small amount of money away.

*Evidence.* Contingency management has shown a positive medium mean effect size in traditional behaviour change treatments [Prendergast et al. 2006]. However, in terms of virtual rewards deployed in DBCIs, evidence for efficacy is mixed, perhaps because they target Type 2 scarce resources for conscious processing of the reward. One short-term (10-day) study found no effect [Zuckerman and Gal-Oz 2014]; while qualitative research both supports [Fritz et al. 2014] and does not support [Munson and Consolvo 2012] the strategy. Adams et al. 2017 found that in a relatively long term intervention (4 months), small, immediate rewards had a greater impact on number of steps than larger, delayed rewards. Evidence for the long-term efficacy of *financial* rewards is also mixed: Volpp et al. 2008 found that although financial incentives achieved weight loss in the short term (3 months), weight loss was not sustained in the longer term (7 months) and the long-term efficacy is unclear. Stawarz et al. 2015 identified a key challenge in delivering rewards to drive the generation of automatic behaviour: positive reinforcement over 4 weeks hindered automaticity, possibly due to reactance. Zedelius et al. 2011 also found evidence that consciously-processed rewards impaired task performance.

*Challenges.* Challenges to implementing an effective reward strategy include issues of monitoring, designing rewards and reward schedules, and theoretical issues. Firstly, accurately monitoring context and behaviour deliver rewards smoothly (so that the user assigns credit from the reward to the correct action [Maia 2009]) is not trivial, as we discuss in Section 6.1. Crucially, the desired behaviour may be extinguished by inaccuracy: both when a given action no longer attracts the previous reward, or when a reward is received despite the appropriate action not occurring [Yin and Knowlton 2006]. Secondly, it is not clear how to apply results from psychology labs to designing rewards and reward schedules for DBCIs. From behaviourism, rewards delivered at certain intervals (interval schedules) promote context-response habit links, while rewards delivered on a given number of responses (ratio schedules) tend to promote action-outcome goal-directed links [Gasbarri et al. 2014; Yin and Knowlton 2006]. Continuous or very frequent rewards can support the acquisition of new behaviours, but develop behaviour that is easily extinguished [Villamarín-Salomón and Brustoloni 2010]. As noted above, consciously-processed rewards may not have the intended effect since deliberative cognitive resources are limited.

Finally, on a theoretical level, a potential problem with punishment or removal of rewards to change behaviour (what behaviourists call *extinction*) is that the underlying associations are not altered: instead participants may simply learn to inhibit the unwanted behaviour in particular contexts [Bouton 2014; Redish et al. 2007].
A common mechanism related to rewards and behaviour in DBCIs is gamification. Gamification extends beyond virtual rewards into engagement. Given that behavioural repetition is crucial to habit formation, engaging users in habit-formation apps is crucial. Nevertheless, many health apps have employed gamification elements without consideration of the underlying behavioural theory [Lister et al. 2014]. Orji et al. 2017 provide evidence that a variety of gamification strategies including “competition, simulation, self-monitoring and feedback, goal setting and suggestion, customization, reward, social comparison, cooperation, punishment and personalization” have different perceived persuasive impacts on different personality types. Further work is required to identify the persuasive impact on behaviour of each gamification strategy. The tendency for DBCI trackers to employ multiple gamification strategies makes it difficult to determine why they succeed or fail. The Basis tracker tried to instil simple habitual behaviours by rewarding repetition [mybasis.com 2015], although it also used a gamification strategy to ‘unlock’ additional behaviour tracking, which may provoke reactance in users who wish to have these locked behaviours tracked at the start. Activity trackers also employ goal-setting techniques, virtual rewards such as progress badges, and enable social sharing of data [Fitbit 2018; Jawbone 2018]. Again, this pick-and-mix strategy approach does not incorporate evidence that different personality types respond differently to different motivational elements [Jia et al. 2016].

Applying virtual rewards and other gamification strategies is also not a cure-all. There is no definitive evidence that either financial or virtual rewards can survive the jump from psychology labs to real-world DBCIs. DBCI designers face an awkward set of choices: what response to reward and how to detect it; what form of reward(s) to use; on what reinforcement schedule; and whether to target conscious or nonconscious processes. Since context-detection is not a solved problem (see Section 6.1), designers also need to consider the adverse impact of wrongly reinforcing behaviour due to technical failures –whether in the wrong context, at the wrong time according to a reinforcement schedule, with the wrong reward, or targeting the wrong Type 1 or Type 2 process.

4.3.2 Self-monitoring (A1, P5).

Outline. Self-monitoring involves using information from self-tracking to form alternative intentions to act [Snyder 1974]. Self-tracking, the capture and presentation of information about an individual’s behaviour, often has a role in revealing information that may be previously unknown to the user [Thaler and Sunstein 2008], such as the number of steps taken each day. It can be helpful to reveal the consequences of automatic Type 1 processes to Type 2 behavioural override mechanisms [Hermansen et al. 2016]. The “self-monitoring and feedback” approach is common in DBCIs [Orji et al. 2017]. Domains include energy usage [Brynjarsdottir et al. 2012], water usage [Kuznetsov and Paulos 2010; Laschke et al. 2011] and activity tracking [Fitbit 2018; Moov 2018], the latter because it is relatively easy to monitor [Ananthnarayan and Siek 2012]. Many activity trackers implement data analysis and reminder engines over and above simple data presentation e.g. Jawbone’s Smart Coach [Jawbone 2018] and Moov’s Fitness Coach [Moov 2018]. With our snacking example, the user could record the amount of unhealthy snacks that they eat in front of the television using a simple self-monitoring app to inform different behavioural decisions.

Evidence. One meta-analysis of 138 interventions found evidence that self-monitoring led to small-to-medium changes in health goal attainment [Harkin et al. 2016]. Self-monitoring weight, i.e. tracking the consequences of undesired eating behaviour, can be an effective long-term strategy in maintaining weight loss for more than 1 year [Butryn et al. 2007; Wing et al. 2006]. Butryn et al. 2007 hypothesize that frequent weighing increases a person’s vigilance over their diet, but also note that reversing small weight gains (made apparent by frequent weighing) may be easier than reversing larger weight gains. Thus, self-monitoring weight enables the activation of smaller
adjustments to diet to intervene early. Kelley et al. 2017 recently provided qualitative evidence that self-tracking can have a role in revealing unhealthy or unexpected behaviour and thus provide motivation for change. Hollis et al. 2015 found some evidence in a 3-week study that simple emotion tracking can augment self-reflection, particularly for reflecting on reasons for failing to stop bad habits.

However, self-monitoring is not a panacea for behaviour change [Epstein et al. 2016; Stawarz et al. 2015]. Evidence shows that use of activity trackers tends to tail off in the longer term. A survey found that more than 50% of US activity owners no longer use their device, with 1/3 stopping using it within 6 months [Ledger and McCaffrey 2014], although the number of participants and survey procedure is unclear. Another study found that even when 17 participants were given $1,000 to purchase smart devices to pursue their goal, almost 80% of devices were abandoned within the first two months [Lazar et al. 2015]. Even when trackers are not abandoned, there is evidence to question their long-term efficacy: a large-scale randomised controlled trial found that using a tracking device alongside self-monitoring of diet and activity resulted in less weight loss compared to a self-monitoring group alone [Jakicic et al. 2016]. Users report abandoning DBCI trackers because of a lack of accuracy, lack of understanding of the mechanics of tracking and difficulties in assessing accuracy [Yang et al. 2015].

Challenges. The mechanism linking self-monitoring with behaviour change is unclear. For example, self-weighing may function as an explicit input to Type 2 processes, affecting deliberative food choices. Alternatively, self-weighing may prime Type 1 processes, triggering nonconscious restraint [Brunner and Siegrist 2012; Pacanowski et al. 2015]. Self-monitoring may be an effective behaviour change technique in the eating domain since the remedy (eat less) is relatively simple; the technique may be less applicable to more complex behaviours and solutions. Further, the design of self-monitoring systems is crucial: the results of self-tracking should be designed so that it requires only limited deliberative cognitive resources to see and interpret results e.g. as implemented in “glanceable” versions [Consolvo et al. 2008a], or minimising effort by using aggregated wellbeing scores [Lin et al. 2012; Meyer et al. 2014].

Self-monitoring is rarely implemented as a stand-alone strategy in DBCI, which makes evaluating its efficacy more difficult. It is often combined with goal setting, goal tracking and goal feedback mechanisms [Consolvo et al. 2008a; Fitbit 2018]. Several DBCI employ self-tracking techniques that require users to track their own behaviour e.g. several habit-specific apps [Way of Life 2018; Xavier 2018], even when the behaviour is one that can be automatically tracked like activity [Chini et al. 2012]. Dual Process Theory and habit theory suggests that self-report of behaviour is unlikely to be accurate, a prediction borne out by empirical data in the activity domain showing substantial differences between self-report and actual sedentary behaviour [Clark et al. 2009; Colbert and Schoeller 2011].

Awareness of the outcomes of Type 1 processes does not necessarily translate into corrective action. A survey of self-trackers found that those who did not also track triggers and context did not have enough information to improve their tracking measures [Choe et al. 2014]. We therefore argue that a key underexploited use for self-tracking in behaviour change is the use of pervasive technology to track trigger cues for unwanted behaviour, rather than simply the behaviour itself.

4.4 Summary

Which strategies, then, hold the most promise to take advantage of this opportunity and create effective DBCI that can change behaviours? Although the HAM bridges the theoretical gap in understanding how both Type 1 and Type 2 processes can impact on behaviour, there is a corresponding empirical gap [Hofmann et al. 2008; Sheeran et al. 2017b]. Specifically, there is a
lack of evidence relating the efficacy of each strategy to alter habits. Table 4 shows an overview of the identified strategies, together with the processes they target, which component of habits they relate to (illustrated with reference to the HAM Figure 1), a summary of the challenges and recent implementations. The challenges are: correct context detection; mixed or missing evidence for a strategy’s efficacy in relation to habit changes; difficulties in translating strategies from psychology labs to in-the-wild DBCIs; and interventions that rely on limited deliberative cognitive resources. Context detection issues and empirical gaps affect most strategies.

Note that although habitual context-response links arise from a Type 1 process, strategies that target them do not necessarily need to be Type 1 or nonconscious themselves. Many of the strategies have both a Type 1 and a Type 2 component. However, given that cognitive resources are limited, we suggest the use of strategies where Type 2 processes are targeted via preparatory training rather than via just-in-time interruptions that are delivered when the user may not have resources to attend to the DBCI. Designers and DBCI research should focus on (a) interventions that aim to automate any conscious Type 2 processes into nonconscious Type 1 processes; (b) strategies where meta-analysis shows evidence of impact on behaviour either through digital or non-digital behaviour change intervention. Based on these criteria, we suggest that implementation intentions and automation of self-control are good candidate strategies for future habit targeting DBCIs.

A caveat to employing implementation intentions is the ongoing challenges in context detection. From Table 4, it is clear that all the suggested strategies, except for providing information, may be hampered by this issue. This may explain why information-providing DBCIs are common. Nevertheless, where cues can be identified, for example in the healthy eating domain, DBCIs should exploit interventions that: do not tax deliberative cognitive resources: alter the context where the costs are not prohibitive; alter the context salience through CBM-Attention techniques; use CBM-Approach and priming. Priming and CBM techniques try to directly target Type 1 processes (Priming targets F1 & P1; CBM-Attention targets F2; CBM-Approach targets P1 in Figure 1). CBM techniques may be used to simultaneously try to lessen the impact of a cue linked to an unwanted response in favour of trying to increase the availability and/or liking of a cue linked to a wanted response. These techniques, by focusing on automating the required response, reduce the likelihood of reactance, and can support habit formation where the automated behaviour occurs in a stable context.

Strategies that target only Type 2 processes in a just-in-time manner (targeting P3 in Figure 1) are unlikely to succeed in changing behaviour with a large habitual element both because they require sparse cognitive resources and because of the challenges in delivering a just-in-time system. Nevertheless, the use of self-control training (targeting Prepare - Type 2) is a candidate for future research as a strategy for both habit-making and habit-breaking because of evidence that it may itself become automatic through appropriate practice, resulting in it emanating from Prepare - Type 1 processes instead, making it more persistent. We need more research into appropriate revalue outcome strategies to form and break habits given the issues we have identified in providing variable rewards in attributable ways.
### Table 4. Strategies, HAM phase and component, and Type 1 / Type 2 targets

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Filter</th>
<th>Prepare</th>
<th>Action</th>
<th>Context detection</th>
<th>Mixed / missing evidence</th>
<th>Translation</th>
<th>Cognitive limits</th>
<th>Recent implementations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alter context</td>
<td>□</td>
<td>□</td>
<td></td>
<td>● ●</td>
<td></td>
<td></td>
<td></td>
<td>[Kalnikaite et al. 2011; Rogers et al. 2010]</td>
</tr>
<tr>
<td>Instinctive primes</td>
<td>■</td>
<td>■</td>
<td></td>
<td>● ●</td>
<td></td>
<td></td>
<td></td>
<td>[Bateson et al. 2006;?]</td>
</tr>
<tr>
<td>Goal priming</td>
<td>■</td>
<td>□</td>
<td>□</td>
<td>● ●</td>
<td></td>
<td></td>
<td></td>
<td>[Chen et al. 2014; Magaraggia et al. 2014; Pinder et al. 2017]</td>
</tr>
<tr>
<td>Alter cue salience</td>
<td>■</td>
<td></td>
<td>□</td>
<td>● ● ●</td>
<td></td>
<td></td>
<td></td>
<td>[Dennis and O’Toole 2014; Enock et al. 2014; Pinder et al. 2016b]</td>
</tr>
<tr>
<td>Train context-response</td>
<td>□</td>
<td>●</td>
<td></td>
<td>● ● ●</td>
<td></td>
<td></td>
<td></td>
<td>[Heitmann et al. 2017; Pinder et al. 2016b; Rabinovitz and Nagar 2015; Wittekind et al. 2015]</td>
</tr>
<tr>
<td>Implementation intentions</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>■</td>
<td></td>
<td></td>
<td></td>
<td>[Breers et al. 2017; Cullen et al. 2016; Pinder et al. 2016c; Veling et al. 2014]</td>
</tr>
<tr>
<td>Provide information</td>
<td>■</td>
<td></td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td>[Comber and Thieme 2013]</td>
</tr>
<tr>
<td>Just-in-time reminders</td>
<td>■</td>
<td></td>
<td></td>
<td>● ● ●</td>
<td>●</td>
<td></td>
<td></td>
<td>[Ding et al. 2016; Klasnja et al. 2015]</td>
</tr>
<tr>
<td>Self-control</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>● ● ●</td>
<td>●</td>
<td></td>
<td></td>
<td>[Blackburne et al. 2016; Cranwell et al. 2014; Lawrence et al. 2015; Miles et al. 2016]</td>
</tr>
<tr>
<td>Self-monitoring</td>
<td>□</td>
<td>●</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td>[Hollis et al. 2015; Kelley et al. 2017; Moov 2018]</td>
</tr>
<tr>
<td>Revalue outcome</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>● ● ●</td>
<td></td>
<td></td>
<td></td>
<td>[Adams et al. 2017; Fitbit 2018; Khot et al. 2015]</td>
</tr>
</tbody>
</table>

- **Direct target**
- **Indirect target**

*Note that self-control also has a Type 1 component that is less affected by cognitive limits than its Type 2 counterpart.*
If the target behaviour is not repeated in a stable context, then there may be a role for just-in-time interventions to ‘remind’ people of their behavioural intentions at opportune moments, if this can be done without attracting reactance. DBCIs can also support users to form intentions in line with the Theory of Planned Behaviour, augmented with Goal Setting Theory, because of the strong empirical evidence for its efficacy in Type 2 situations. These interventions should be accompanied by priming and self-control strategies as outlined above. However, habitual behaviour should not be expected to result because of the absence of a stable context. The COM-B model and the Behaviour Change Wheel can also be used to ensure wide consideration of other non-habitual strategy options.

We address the issue of context detection in depth in Section 6, alongside a set of general difficulties in designing and evaluating DBCIs. First, in the next section, we bring together a set of design principles in applying dual process interventions for habits.

5 DESIGN PRINCIPLES

In applying the HAM to DBCIs targeting habits, we have identified several common design principles. These principles clarify the lessons learned and provide a starting point for intervention design. They also point to further challenges, which we expand on in Section 6. Some of these principles, such as understanding target behaviours, context, tailoring and ethics, are common to all DBCIs. However, there are specific issues within these principles that habit-targeting DBCIs need to consider.

5.1 Understand and simplify target behaviour and context

As with all DBCIs, the starting point is to clearly identify the target response [Michie et al. 2014a]. For DBCIs seeking to form habits, there is a particular advantage to simplifying the target behaviour as much as possible. The simpler the target response, the faster automaticity can occur [Lally and Gardner 2013; Wood and Neal 2007]. Since we define habits as behavioural impulses triggered by contextual cues, understanding the context is crucial. These cues may include those internal to the user, e.g. mood or physiological states like hunger, as well as more obvious external cues such as location and time. Habit-forming DBCIs should select the smallest possible set of salient cues to form a ‘stable context’ as a habit trigger, since simpler context causation models support faster habit formation than complex ones [FitzGerald et al. 2014]. DBCIs should avoid re-using contextual cues already in use in context-response links, because cues associated with multiple responses are more likely to cause response conflict, triggering arbitration and conscious Type 2 resources, and thus hinder habit formation [Wood and Neal 2007]. Habit-breaking DBCIs need to isolate the particular set of context cues that trigger an unwanted response. For example, when considering an intervention to stop snacking, the DBCI designer needs to assess when, where and why the unwanted behaviour occurs. If they determine that the snacking occurs at home under conditions of being distracted by the television, the DBCI can isolate specific contextual cues related to the television to provide the basis of an intervention, for example within an implementation intention. It is important to note that this level of rich context-detection is not trivial: Section 6.1 discusses the specific challenges such detection.

Context cues may also include responses themselves via chaining. Chaining is the linking of a new behaviour to an existing habit, which can support habit formation [Judah et al. 2012]. Within the HAM, successful chaining means an existing behavioural response acts as a cue input to trigger the new required response. For example, to promote healthy snacking in the presence of habitual TV watching, the user should be encouraged to make healthy snacks available next to the TV watching area.
5.2 Type 1 / Type 2 tailoring

It is a general DBCI principle that interventions should adapt to individual users [Ijsselsteijn et al. 2006; Orji et al. 2017; Ranfelt et al. 2009]. The HAM requires a specific form of tailoring because individuals vary in relative influence of Type 1 and Type 2 behaviours [Sladek et al. 2006]. For example, one individual may be more impulsive or susceptible to temptations such as the presence of consciously unwanted snacks than others. In these circumstances, a DBCI may need to intervene earlier in the unwanted behaviour process, e.g. by altering the context to try to prevent a user from buying unwanted snacks in the first place to remove the tempting cue. Individual users at different points in the dynamic process of habit formation or breaking may also require different sorts of intervention. For example, in the early stages of habit formation via implementation intentions, a user may need a higher level of support via intention rehearsal and reminders. These reminders should decrease over time as behavioural automaticity emerges to avoid technological dependence [Renfree et al. 2016] and possible reactance, as we discuss in principles Section 5.4 and Section 5.6 below.

5.3 Design for Type 1 and Type 2 congruence

Since the HAM shows that behaviour may result from simultaneous influences of Type 1 and Type 2 processes, the most effective habit-focused interventions are likely to be those that arise from influencing Type 1 and Type 2 processes in congruence. We have argued above that a likely failure of many interventions is the focus on Type 2 processes, undermined by incongruent Type 1 default processes. We caution against a corresponding myopic focus on Type 1 processes only without ensuring that the behaviour change participant is at least also consciously motivated to change their behaviour in a given direction. For example, priming can only be successful where a person already has relevant cognitive constructs motivated towards the given behaviour [Strahan et al. 2002]. If, in our unhealthy snacking example, the person has no interest or motivation in changing their snacking to a healthier alternative, priming them to switch snacks will not be successful. Congruent DBCIs that focus on Type 1 and 2 processes in unison would provide support for user behaviour change regardless of levels of user attention, deliberative resources and individual differences in Type 1 / Type 2 dominance.

5.4 Design for persistence

Habitual behaviour change is a long-term process. Automatic behaviour only emerges over time, so habit-focused DBCIs need to be viable over the longer term. One approach to establishing more persistent interventions is opportunistic training or incidental interaction [Ding et al. 2016; Dix 2002], fitting into and/or appropriating existing user behaviours. This design fits well with both COM-B and Fogg Behavior Model opportunity constructs. DBCIs should also follow the principle of synergy not substitution [Sellen and Whittaker 2010]. The aim should be to develop interventions that complement, and indeed augment [Rogers 2006], how our brains work to leverage the habit system, rather than leaving users dependent on their machines to substitute for unwanted brain processes. In particular, using just-in-time reminders based on technology could introduce a dependence on that technology, if it becomes part of the context-response link. The response is then more fragile, and will not occur if the technology context is abandoned [Renfree et al. 2016]. Technology abandonment is a known issue in the fitness tracker domain [Goodyear et al. 2017; Yang et al. 2015]. DBCIs instead should use training paradigms such as cognitive bias modification to alter default reactions to cues, create dependence on stable contextual cues, for example via Implementation Intentions, or co-opt existing behaviours such as unlock gestures as training methods to deliver cognitive bias modification or priming interventions [Pinder et al. 2016b, 2017].
For example, a smartphone DBCI might present a serious game on every unlock that requires a user to repeatedly pair an image of a television with an image of the healthy snack the user wishes to consume while watching television. Note that the requirement to tailor to Type 1/Type 2 differences over time may also require the DBCI to determine when to self-destruct once it is no longer necessary to the performance of the target behaviour. We explore the evaluation implications of a longer time frame in Section 6.2.1.

5.5 Design for multiple points of intervention

DBCIs should look beyond the main technology platform of the DBCI to determine whether other interventions (both digital and analogue) can be included, given evidence that multiple component interventions (e.g. apps plus face-to-face counselling) can outperform apps alone [Schoeppe et al. 2016]. For example, a DBCI could instruct a user to print photos to act as persistent primes to consume a particular healthy snack in a particular location, in addition to supporting them to form and rehearse implementation intentions to consume that snack in the given location.

5.6 Design for reactance

Interventions that attempt to change a user’s behaviour threaten a user’s autonomy [Roubroeks et al. 2011], and may provoke reactance. Reactance occurs when users respond to a perceived restriction in behavioural freedom by trying to regain that freedom [Brehm 2009]. It may cause a user to disengage from a DBCI entirely or they may take steps to limit perceived infringements of freedom, e.g. by disabling notifications. Reactance may be triggered by the content of a persuasive message [Roubroeks et al. 2011] or the timing of its delivery. Inappropriate suggestions, lack of personalisation and monitoring errors may damage the credibility of the system [Ding et al. 2016; Segerståhl et al. 2010] and thus trigger reactance. Shame (e.g. from failing to achieve behavioural goals) or “excessive competition” may alienate users [Consolvo et al. 2009a].

DBCI designers should therefore ensure that just-in-time behavioural directions (i.e. those that directly threaten users’ behavioural autonomy) should only be delivered when the system has confidence that the timing is appropriate. For example, to avoid issues of inappropriate timing, one strand of Ding et al.’s intervention suggested that users continue with existing walking behaviour, rather than simply directing them to start walking [Ding et al. 2016]. Where it is not possible to make such judgements of appropriateness, DBCI designers may instead choose to use a low-reactance alternative, e.g. priming or implementation intentions. For example, instead of issuing directives to start eating a particular healthy snack, a DBCI might support a user to form an implementation intention linked to snacking moments, such as ‘if I am watching TV, then I will eat the healthy snack’. We explore the challenge of reactance further in Section 6.4.2.

5.7 Design ethically

All DBCI designers have a moral duty to ensure that their interventions are ethical. This is a particular consideration where techniques are used to target Type 1 nonconscious processes, e.g. subliminal priming and cognitive bias modification. There is a certain amount of ‘moral panic’ about subliminal priming, despite evidence that the technique is only effective where the priming targets a goal the user is already motivated towards [Strahan et al. 2002]. As shown in the HAM, priming and other Type 1 targeting techniques can only activate pre-existing associations within memory.

Given these concerns, it is important for users to provide active and informed consent to their DBCI based on full disclosure of how the DBCI works. For example, a DBCI that uses subliminal priming should be clear both about the prime used and the intention with which it is presented; a cognitive bias modification intervention should be clear about the end-goal of the training. The
ultimate aim of habit-focused DBCIs is to build ‘good’ habits, persistent behaviours that people are not fully aware of enacting. Therefore, it is even more critical that people are able to give fully informed consent about the implications of habit-targeting DBCIs. Ultimately, users should be in control of their DBCIs, rather than the other way round [Rogers 2006]. They need to understand what the DBCI is doing and why, and have some ability to measure its impact and to remove and/or reverse its effects if required. This principle is often violated with non-digital behaviour change techniques such as tobacco packaging warning messages [Peters et al. 2013]. We address the ethical challenges further in Section 6.3.

6 CHALLENGES IN DELIVERING HABIT-FOCUSED DBCIS

This section brings together the challenges we have identified in designing and evaluating habit-focused DBCIs. The intention is to highlight areas of future research, and to assess what is possible given current technology.

6.1 Context detection

There are a number of technical challenges to implementing just-in-time context- and/or response-targeting strategy such as “anticipatory” interventions [Pejovic and Musolesi 2015]. To detect and intervene in habitual behaviour, the technology needs to:

   a) accurately detect and/or predict a response;
   b) accurately monitor the context to determine whether each cue(s) is prompting the unwanted response (for habit-breaking) or suitable candidate contextual cue(s) on which to attach a wanted response (for habit-making); and
   c) direct the user to avoid the context for the unwanted response (habit-breaking) or alert them to the context for the wanted response (for habit-making) at the right time, in such a way that the user complies and without causing irritation; and
   d) optionally also provide rewards for desired responses and punishments for undesired responses consistently to the best schedule for the given context.

All of these steps are not solved problems. For steps (a) and (b), problems of accurate behaviour and context-detection are well known within ubicomp [Bettini et al. 2010; Rogers 2006]. The accuracy of response tracking also varies with domains: capturing smoking or eating is more difficult than using smartphone and/or activity tracker accelerometers to capture sedentary behaviour. Research is ongoing to capture non self-report data in both the smoking e.g. [Scholl et al. 2013] and eating domains e.g. MyBehavior 2.0’s crowd-sourcing of calorie information of food pictures [Rabbi et al. 2015]. Yet even tracking activity via accelerometers is not trivial: proper evaluation metrics for activity recognition are not being sufficiently considered [Lukowicz et al. 2012].

A context cue may include a physical location, a particular time period, the co-presence of others or cognitive constructs such as mood [Ji and Wood 2007] or some representation of a goal. The first two are relatively easy to track. Tracking the presence of other people depends on them carrying (and having configured in a certain way) identifying technology such as Bluetooth or users accepting more invasive emerging technology such as skin-mounted RFID tags [Ziai and Batchelor 2011]. Tracking cues that are internal to the user like mood and emotions is much more difficult, despite advances in physiological computing [Fairclough 2009; Hernandez et al. 2015], the increasing use of Experience Sampling Methods [Aharony et al. 2011; Rachuri et al. 2010] and behavioural causation analysis on smartphones [Pejovic and Musolesi 2015].

It is also difficult to predetermine which cues from a given context pass through perception and implicit filters (see Figure 1) for a given individual. A target context suitable to be selected as a trigger for habit formation needs to be “sufficiently salient in daily life that it is encountered and detected
frequently and consistently” [Gardner et al. 2012b]. It is not clear what level of detection either in Type 1 or Type 2 processes must be achieved for a context to become a trigger, nor how technology might determine which individual cues within a context have been ‘noticed’. Nonetheless, it is an open question the extent to which users can tolerate inaccuracy. Minor accuracy issues are likely to pass unnoticed and may still provide the user with useful information about their behaviour, context or mood.

Given these difficulties, it is perhaps not surprising that relatively few DBCIs use context analysis [Honka et al. 2011; Stawarz et al. 2015]. Implementations are starting to emerge [Lee et al. 2017; Naughton et al. 2014, Rabbi et al. 2015], albeit ones partially reliant on self-report. Lee et al. 2017 built a system to support users to create their own context-aware if-then plans to deal with sleep problems (e.g. to remove technology from a certain location at a certain time), while the MyBehavior system aims to provide context-aware strategic behaviour change suggestions (e.g. to continue with an existing walk) [Rabbi et al. 2015].

Our HAM framework allows internal conditions such as mood to form part of the Context or set of cues that can trigger habitual behaviour. However, few DBCIs use automated mood tracking, although (i) some trackers use other physiological markers e.g. respiration as an indicator of emotion [Spire 2018] and electrical brain activity as an indicator of calmness [Muse 2018], and (ii) self-report of emotive states is starting to appear, albeit largely as a peripheral characteristic [Hollis et al. 2015]. Perhaps the most promising avenue of research is the use of machine learning techniques on smartphones to estimate context, behaviour, and internal cognitions such as mood [Burns et al. 2011]. Nevertheless, the research is still at a preliminary stage and training of mood involves relying on self-report. As we outline in Section 6.2.3, self-report is not reliable in all circumstances.

For step (c), directing and interrupting users can trigger reactance, where users reassert their own control over their behaviour in response to perceived threats. As we discuss further in Section 6.4.2, this is a key challenge. In any case, directing user behaviour via just-in-time prompts is a difficult strategy. Firstly, the prompts may be delivered as persuasive messages via notifications on wearable technology (e.g. smartphones, smartwatches). However, research into attention shown to interruptions on smartphones shows that users may take more than 10 minutes to respond to a notification on average [Pejovic and Musolesi 2014], which implies that just-in-time interruptions on such devices are likely to fail. Secondly, regardless of prompt format, interruptions on any platform may also be counter-productive: interruptions deplete self-control resources [Freeman and Muraven 2010], and can cause disruption even if they contain important or useful information [Mehrotra et al. 2016].

To counteract the interruption problem, indirect forms of prompting could be used as a proxy for notifications, for example vibration and/or audible triggers as in SitCoach [Dantzig et al. 2012], although other research found that most users ignored vibration interruptions [Hirano et al. 2013]. Alternatively, interventions can use one of the emerging interruptibility prediction models [Kapoor and Horvitz 2008; Mehrotra et al. 2015], where interruptions are only issued where the potential cost of the user not heeding the interruption outweighs the potential cost of interrupting now and getting the timing wrong.

Future research to address ‘ideal’ interruptions should include analysis of how to support users to construct appropriate interruption rules, potentially from large amounts of contextual data. Several researchers exploring tailored DBCIs have found that users tend to limit their if-then rules to simple rules in small numbers, due to issues in understanding more complex sets of rules [Lee et al. 2017; Pinder et al. 2016c].

For step (d), as noted in Section 4.3.1, establishing the optimal schedules of reinforcement for rewards/punishments for a given response is also difficult.
6.2 Evaluation

Behaviour change interventions are difficult to evaluate [Intille 2004; Klasnja et al. 2011] for three main reasons: they need to be monitored in the long term, use an appropriate experiment design, and use an appropriate measure of behaviour.

6.2.1 Long-term monitoring. In the Design for Persistence design principle (Section 5.4), we noted that habits take time to change. Thus long-term evaluation is crucial for DBCIs that target habits. Lally et al.’s study [2010] of a once-a-day repeated behaviour showed that it took between 18-254 days to plateau into automaticity, with a median time of 66 days and substantial variation at an individual level. A study of exercise habitual behaviour found that to establish a new exercise habit, new gym-goers had to exercise at least four times per week for around six weeks [Kaushal and Rhodes 2015]. However, most HCI research methods do not lend themselves well to determining whether a given intervention has resulted in a permanent change in behaviour [Fogg and Hreha 2010; Klasnja et al. 2011], yet post-intervention monitoring is crucial to determine whether changes in habitual behaviour are stable. The gold standard is the sort of Randomised Controlled Trial used in medical research, but large-scale long-term trials are difficult to establish in HCI [Hekler et al. 2013; Klasnja et al. 2011]. Alongside this, there is the familiar issue of studies using small numbers of participants, no control groups and not reporting behaviour change results.

Empirical evidence demonstrates the difficulties: a meta-review of internet-based DBCIs rejected 26% of 549 possible studies because they did not use a measure of behaviour, and 15% because they did not use a control group [Webb et al. 2010a], while Froehlich et al. [2010] analysed 8 in-the-wild studies in the energy domain and found an average of 11 participants lasting 1-4 weeks, where no studies had a control group and only 4 reported behaviour change data. These findings are supported by Brynjarsdottir et al. [2012], who reviewed persuasive sustainability research and found almost half of the papers had no user evaluation, while field studies typically lasted 3-4 weeks with fewer than 10 participants; and Hamari et al. [2014], whose review of 95 DBCIs found that experiment length was “very short in most cases”, and sample size median was 26 (where reported). On an individual study level, two highly cited trials in the health behaviour change domain, UbiFit [Consolvo et al. 2008b] and Fish’n’Steps [Lin et al. 2006] had small numbers of participants (12 and 19 respectively), although Fish’n’Steps was tested over 14 weeks (including pre- and post-baselines) to UbiFit’s 3 weeks (no baseline).

The short-term nature of DBCI testing may also skew empirical support for theoretical constructs. Sheeran et al. provide evidence that initial experience of a new behaviour strengthens the intention-future behaviour link, while substantial experience attenuates the link due to habit formation [Sheeran et al. 2017a]. Thus short-term studies introducing participants to novel behaviours may provide evidence of a strong intention-behaviour link that does not persist over time. DBCIs that target habitual behaviour also face more specific issues of evaluation in monitoring and measuring users, which we outline below.

6.2.2 Experimental Design. There are several tensions in establishing the correct experiment design to explore habitual behaviour. The first tension, in line with evaluation issues, is in the length of study. As noted above, short term evaluations may miss key habitual changes. However, Kahneman suggests that short experiments with between-participants designs and little information about the purpose of the experiment are preferable to explore Type 1 processes [Kahneman 2003]. This is because longer designs with repeated measures “encourage the participants to search for consistent strategies to deal with the task” [Kahneman 2003].

---

9335 citations in the ACM Digital Library, 990 in Google Scholar on 22 March 2018
10247 citations in the ACM Digital Library, 723 in Google Scholar on on 22 March 2018
The second tension is that many experimental designs are inadequate to capture intra-individual variations in behaviour [Gardner 2015]. This is particularly key given evidence of individual variation in the relative impact of Type 1 vs Type 2 processes in determining behaviour [Sladek et al. 2006]. Appropriate study designs are those that: can use mixed-level modelling to capture within-person variation [Jaeger 2008]; ‘N-of-1’ experimental designs, which can reveal how behaviours change over time [Klasnja et al. 2017; Lillie et al. 2011; McDonald et al. 2017]; or micro-randomized trials examining the impact of individual behaviour change components on proximal outcomes [Klasnja et al. 2017]. The latter is a variation on a sequential factorial design: participants are randomly assigned different interventions at times judged to hold potential for effective intervention [Klasnja et al. 2017]. It holds promise as a way forward in the pick-and-mix DBCI design context to judge the efficacy of different intervention elements.

6.2.3 Appropriate measure of habits.

Measuring behaviour. Directly measuring actual behaviour is the gold standard of a behaviour change DBCI, in line with a behaviourist standpoint, but is difficult to achieve in practice [Klasnja et al. 2011]. Few health game interventions are measured beyond changes in behavioural intention [Hwang and Mamykina 2017]. Even when behaviour is measured, e.g. [Cambo et al. 2017], it is important to also measure baseline pre-intervention and post-intervention behaviour to determine whether (a) behaviour has changed and (b) whether that change is sustained in the longer term as outlined above.

Self-report. Habit measurement is therefore dominated by self-report of behaviour [Gardner 2015]. The two main self-report measures for habits are Verplanken & Orbell’s twelve-item Self-Report Habit Index (SRHI, 2003) and the Self-Report Behavioral Automaticity Index (SRBAI, [Gardner et al. 2012a]). The SRHI is designed to gauge habit strength and determine the role of habit without measuring behavioural frequency. The SRBAI is a subscale of the SRHI to measure automaticity more succinctly. Verplanken & Orbell 2003 found a strong correlation over four studies between the scores in the SRHI and habit frequency, and note that the use of behavioural frequency as a measure of habit is itself only a proxy for a true measure of habit strength.

Although these two measures have been used widely (e.g. SRHI [Gardner et al. 2011], SRBAI [Kaushal and Rhodes 2015]), there are concerns about their validity [Nilsen et al. 2012]. Since we define a habit as an impulse for a response triggered by a particular context, arising in Type 1 processes, self-report measures requiring Type 2 self-reflection on behaviour are insufficient. The indices measure the experience of habitual behaviour rather than the underlying processes, and a person’s recollection of automaticity is unreliable as a habit measure [Hagger et al. 2015]. Further, the SRHI does not capture the impact of cues, and is limited to capturing the consequence of enacted habits [Sniehotta and Presseau 2012].

Nevertheless, self-report should not be disregarded entirely. There is some evidence that self-report can be helpful in triggering enhanced reflection around possible causes of problematic behaviour, particularly when emotions are self-reported [Hollis et al. 2015].

Self-report alternatives. Association tests, measuring response time speed and accuracy of context-response links, should be considered as measures of habit as alternatives to self-report [Aarts et al. 1998; Gardner 2015]. There is evidence that these techniques can provide better predictions of ongoing behaviour than those arising from self-reports [Mccusker 2001]. There are a broad range of techniques of evaluating the Type 1 activity are appropriate for DBCIs [Wiers et al. 2013], each one providing an indirect measure of Type 1 activity by providing a reaction time measure to given stimuli. Three commonly used techniques, which may also be appropriate for DBCIs, are Stroop tests, Implicit Association Tests and Dot Probe Tests. Stroop tests (including emotional Stroop
variants), which measure cognitive control and emotional attachment [Williams et al. 1996] are used to measure the relative control of Type 1 vs Type 2 processes in response to a particular cue: it is assumed that the more salient a cue is to Type 1 processes, the more difficult it is for Type 2 to override it, and therefore a longer reaction time. Implicit Association Tests [Greenwald et al. 1998] measure implicit evaluations of concepts and can therefore indicate where a particular action impulse might appear on the HAM’s Potential Response stack (P6 in Figure 1). Dot Probe Tests [MacLeod et al. 1986] can both train and estimate automatic associations of cognitive constructs represented by words and/or images and can thus indicate implicit attentional bias at the Filter stage of the HAM (Figure 1).

However, there is ongoing debate about the validity of associative measures of Type 1 activity including habits [Hagger et al. 2015; Stacy and Wiers 2010]. Furthermore, researchers need to be careful of practice effects [Greenwald and Nosek 2001], implying the need for careful experiment design (e.g. using the latest scoring algorithms) and a control group.

**Multiple measures.** While it remains unclear which is the best method to measure habits, DBCIs should use multiple points of measurement. Measures should include both a measure of habitual behaviour and a measure of the habitual impulse. Until activity recognition algorithms are sufficient to capture behaviour, we suggest using the SRBAI or other validated self-report measure of the target behaviour, together with at least one associative measurement. Health psychologists are sceptical about performing associative tests in the wild [Gardner 2015; Hollands et al. 2016]. However, we suggest DBCIs have a unique advantage in being able to measure context-response associations in situ. Home-based association tests can be a better predictor of behaviour than lab-based tests [Houben and Wiers 2008]. More research is required into how best to translate lab-dependent desktop measurement techniques into DBCIs.

**Not measuring behaviour.** Jepson et al.’s meta-review 2006 of health behaviour change interventions identified an ongoing issue with conflation between knowledge, attitudes and behaviour. The issue also affects DBCI research. For example, a study of interactive shopping mannequins measured *perceived* length of stay at a shop window as an outcome, rather than measuring *actual* length of stay [Reitberger et al. 2009], while research into changing exercise behaviour using a heart rate monitor [Harjumaa et al. 2009] does not report the results of quantitative measures, instead reporting user attitudes and experiences. A focus on explicit attitudes as a determinant of behaviour also arises from a theoretical basis. For example, the use of the Theory of Planned behaviour implicitly accepts a link between attitudes and behaviour.

**What to report.** There is a lack of consensus on how to report DBCIs so that findings are generalisable. As with other behaviour change research, there is a dearth of common reporting standards [Webb et al. 2010b]. Multiple efforts are being made to deliver ontologies of behaviour change interventions to facilitate reporting (e.g. [Klasnja et al. 2017; Larsen et al. 2017; Michie et al. 2013a]). Once a consensus emerges, these should be followed. In the meantime, research must clearly report the intervention, including: theoretical basis; specific behaviour change technique(s); domain; behavioural context(s); mode of delivery (including specific technology); type of behaviour being targeted including the extent to which it is habitual; for HAM interventions, which process (Type 1 vs Type 2) is being targeted; together with transparent statistical reporting [Kay et al. 2016]. We also need more information about participants given the heterogeneity of intervention efficacy [Klasnja et al. 2017; Orji et al. 2017].
6.3 Ethics and privacy
We see much potential for DBCI technology to empower people to take control of their lives and be more proactive about their daily practices [Rogers 2006], particularly by helping them to break undesired habits and create desired habits. However, there are several ethical concerns that need to be considered in order to comply with the design ethically design principle (Section 5.7).

6.3.1 Who is in control? There is much potential power in enabling people to use Type 1-targeting DBCIs to alter their associative memories. However, there are also worrying opportunities for governments and corporations to covertly impose their own agendas on people via these mechanisms [Pinder 2017]. As noted in Section 5.7 above, controversy is particularly high over interventions that use subliminal priming techniques.

6.3.2 Full disclosure and safety. DBCI designers using nonconscious Type 1 methods (e.g. instinctive paths, subliminal priming) need to be explicit about how their intervention works. It is not clear how much disclosure is needed for ethical and safety reasons, and when it may start impacting the effectiveness of the intervention. For example, do we inform users that a smartphone subliminal priming intervention will flash ‘activation words’ on their display, or do we inform them of exactly which words are used and when, or even allow them to play back the prime in slow-motion? We argue that the benefits from an ethical perspective of informing users of how the DBCI works outweigh the drawbacks of it being less effective or ineffective. DBCIs that are only effective when people are (partially) unaware of how they work should be avoided.

A particular issue regarding disclosure is safety. Triggering fear-based flight responses as used in Zombies, Run! [zombiesrungame 2018] in inappropriate contexts (e.g. whilst crossing the road) is clearly problematic. Safety is also an issue when delivering disruptive habit interventions to activate Type 2 processes. Type 2 processes have limited capacity, so disrupting a user when they are already using Type 2 processes in another potentially dangerous task (e.g. driving) is problematic.

Who bears responsibility for the effects of tools to change behaviour is an open question: is the end-user or the system designer responsible [Verbeek 2009]? This issue is particularly apt for configurable systems: for example, what if a user with an eating disorder altered a cue-valence-altering system to devalue all foods instead of just unhealthy foods [Pinder 2017]?

6.3.3 Privacy. Alongside ethics, privacy is an issue with DBCIs, particularly where systems may be disclosing automatically-sensed information to third parties, e.g. for off-device computation. The fusion of real-world sensing with social networks also raises privacy concerns where users have limited control over how they are represented to their social network contacts [Efstratiou et al. 2012]. Again, system designers should be explicit about what data is shared and why, and allow users to opt out.

6.4 Cyberpsychology of interacting with DBCIs
There is a research gap in understanding the cyberpsychology [Norman 2008] of DBCIs that target habits. For example, how different are the reactions to interventions targeting Type 1 versus Type 2 processes? How do these reactions vary between individuals? How does the context in which the habit-targeting technology persuades (e.g. whether it is public or private) impact on efficacy [Mylonopoulou and Isomursu 2016]?

Pervasive technology gives us a great deal of power to support people to change their behaviours by targeting both Type 1 and Type processes, but we need to understand more about how these interventions are perceived. Two key considerations are agency (who is perceived to be doing the persuading) and reactance.
6.4.1 Agency. According to the “Computer As Social Actor” (CASA, [Reeves and Nass 1996]) paradigm, users tend to ascribe social characteristics to even minimal computer interfaces. If this holds true for DBCIs, then they can take advantage of extensive human-human persuasion research [Cialdini 2001], for example by leveraging social priming or contagion [Christakis and Fowler 2013]. However, it is not clear to what extent pervasive systems, particularly unobtrusive ones, may either exploit or avoid this social assumption.

6.4.2 Reactance. As noted in Section 5.6, reactance is a design issue for all DBCIs. It is an open research question to what extent and in what circumstances DBCIs targeting Type 1 processes can trigger reactance. Subliminal pervasive DBCI techniques may avoid reactance, but this has yet to be shown empirically. An unobtrusive DBCI targeting Type 1 processes may enable users to feel free from direct observation by ‘official’ sources and therefore reduce the likelihood of reactance. Users may therefore view unobtrusive DBCIs as potentially more acceptable and effective than externally-imposed solutions. A recent trend towards enabling DBCI participants to choose their own pick-and-mix interventions from a selection of behaviour change techniques [Lee et al. 2017; Meinlschmidt et al. 2016] may prove to be key in avoiding reactance if it makes users feel more in control of their DBCI.

6.5 Theoretical gaps in the HAM
All three of the HAM’s underlying theories – Dual Process Theory, modern habit theory and Goal Setting theory – are active areas of research. Under Dual Process Theory, research is ongoing into determining which process, Type 1 or Type 2, predominates under what conditions [Sladek et al. 2006]. This contrasts with theory development in the health behaviour change domain, which has been dominated by expanding constructs rather than isolating the contexts in which those constructs have most impact on behaviour [Sheeran et al. 2017b]. As outlined in Section 2.7.2, modern habit theory needs to be subjected to more rigorous empirical testing and a more rigorous analysis of the mechanisms underlying formation and automaticity. Likewise, as noted in Section 2.7.3, Goal Setting Theory is being updated to incorporate Type 1 processes [Latham et al. 2017]. In terms of our HAM, further research is also required into effective combinations of the identified habit-changing techniques [Harjumaa et al. 2009], to establish the parameters of effectiveness of any given strategy [Michie and Johnston 2012; Peters et al. 2015], including determining the most effective technology platform on which to deliver it.

Where a behaviour targeted by a DBCI is indivisibly complex - for example some exercise behaviours that involve equipment and/or travel - several researchers have suggested that instigation habits should be distinguished from execution habits [Gardner 2015; Kaushal et al. 2017; Phillips and Gardner 2016].

Instigation habits are the generation of a behavioural impulse for a given behaviour from cues; execution habits are the carrying out of the sequence of sub-actions that make up the cued behaviour. Although our HAM accommodates this distinction (initiation occurs during the filter/prepare stage; execution during the act stage), further research is required to determine the appropriate interventions for each and their relative impacts on behaviour.

6.6 Summary
Despite our optimism that DBCIs targeting habits can achieve long-term behaviour change, several key challenges remain. Context is crucial, but highly accurate cue, prime and behaviour identification are arguably still unsolved ubicomp problems. This makes just-in-time interventions difficult to implement accurately. All DBCIs, but particularly inaccurate just-in-time interventions, risk both alienating users via reactance and inadvertently achieving ironic effects. Interventions
face difficulties in selecting both an appropriate experimental design to test their efficacy, and in selecting an appropriate habit measurement mechanism.

As with any intervention to change behaviour, DBCI designers must consider the ethical implications of both what they promise and what they deliver. Ethics is also an issue where the cyberpsychology of habit-focused interventions is little-understood. Finally, although we have explicitly given our framework a solid theoretical grounding, theoretical gaps remain, particularly in establishing the boundaries of which set of processes predominates in which contexts.

7 FUTURE RESEARCH DIRECTIONS & OPPORTUNITIES

Table 5 shows the habit-altering strategies and the corresponding open research questions derived from the HAM, our overview of the strategies and our design principles. Note from Table 4 only one strategy, implementation intentions, currently has strong evidence to support its use in habit-changing interventions, and that all strategies bar the provision of information face the challenge of imprecise context detection (Section 6.1).

The empirical gap is partly due to the pick-and-mix nature of the research. Pick-and-mix is problematic where multiple strategies are used without using an experiment design that can differentiate their impact (see Section 6.2.2). However, pick-and-mix can be a good intervention where there is evidence that the individual elements are effective. In the broader behaviour change research sphere, multiple strategies tend to be more successful than single interventions [Michie et al. 2016]. The HAM suggests that the most effective approach is likely to be a combination of Type 1 and Type 2 strategies at multiple HAM levels, and research into the interplay of Type 1 and Type 2 processes as determinants of behaviour change is increasing in health psychology [Hollands et al. 2016]. Carels et al. 2014 recently employed a habit-focused weight loss intervention using multiple strategies including breaking unwanted habits, forming new healthy habits, increasing exposure to healthy cues and automatic goal priming. They found a medium to large effect of the intervention at a 6-month follow up compared to an alternative. We suggest a similar approach for DBCIs to identify multiple strategies to employ.

7.1 Opportunities

We see great value in continuing to research how to use context-aware pervasive technology to support highly individualised DBCIs that can deliver a mix of interventions. The focus of research should be on two elements: firstly to use DBCIs to find out which features of HAM are active at different points for different behaviours for a given individual; secondly to exploit this knowledge in DBCIs that allow that individual to pick-and-mix their own interventions.

7.1.1 Trigger hunters. Type 1 context-response associations are not easily available to introspection. Pervasive technology could play a key role in uncovering these associations to discover which contextual features (cues) act as triggers for a particular unwanted response. This may enable us to address the challenge of fluid causal influences affecting both Type 1 and Type 2 systems in the wild [Michie et al. 2013b]. Once uncovered, people can avoid, approach and/or retrain their trigger cues accordingly. We need technology with richer contextual awareness—including mood-detection—to identify both existing cue(s) that trigger unwanted behaviour and candidate cue(s) that can be used as anchors to associate with wanted behaviours within implementation intentions. Advances in context-cue detection and behaviour detection, perhaps driven by machine learning techniques [Banovic et al. 2017], will broaden our understanding of what sorts of cues act as response triggers for different types of behaviours, which can feed into our models of what sorts of cues should be used as habit triggers.
Table 5. Habit altering strategies and research questions

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Key research questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>General HAM application</td>
<td>What is the most effective combination of Type 1 and Type 2 strategies? How should DBCIs adapt as behaviour becomes more automatic? How best to use emotion as a behaviour change strategy?</td>
</tr>
<tr>
<td>Alter context</td>
<td>What advantages does ambient persuasive technology confer over priming on mobile device?</td>
</tr>
<tr>
<td>Priming</td>
<td>What is the best platform/format/schedule on which to deliver these primes? What are the instinctive primes for different contexts and different behaviours? How can we best instil and automate goals?</td>
</tr>
<tr>
<td>Alter cue salience</td>
<td>What is the most effective training strategy for altering attention and approach biases via technology-mediated CBM: opportunistic training or session-based?</td>
</tr>
<tr>
<td>Train context-response</td>
<td></td>
</tr>
<tr>
<td>Implementation intentions</td>
<td>How can DBCIs best support implementation intention formation? What is the most effective way for DBCIs to exploit their additional rehearsal and reminder opportunities?</td>
</tr>
<tr>
<td>Provide information</td>
<td>What is the best method of delivering just-in-time reminders to avoid reactance and avoid overloading Type 2 cognitive limits? Who should set the goals, the DBCI or the user?</td>
</tr>
<tr>
<td>Self-control</td>
<td>What are the determinants of automating self-control?</td>
</tr>
<tr>
<td>Self-tracking</td>
<td>Which type of process does self-tracking operate on, and how can we automate it? How best to track cue triggers, including emotions?</td>
</tr>
<tr>
<td>Revalue outcome</td>
<td>How best to deliver rewards and reward schedules? Do virtual rewards work? Do implicit rewards work?</td>
</tr>
</tbody>
</table>

7.1.2 Self pick-and-mix. A key area of research is in enabling individuals to vary their interventions according to their preferences [Ijsselsteijn et al. 2006; Lee et al. 2017, 2011; Ranfelt et al. 2009], personality traits [Orji et al. 2017] and different digital devices. For example, DBCIs could enable people to use their own images of real-life problematic cues in CBM-Attention interventions. We foresee a crucial role for such DBCIs to help solve the fundamental variability in human behaviours, motivations and contexts [Rogers 2006]. The crucial design points for future DBCIs are that research must be theory-driven, and use experimental design and reporting standards that allow generalisability of results, as discussed in Section 6.2.2 and Section 6.2.3.
8 CONCLUSIONS

We have presented a set of strategic approaches which hold promise for designing DBCIs that alter habitual behaviour to generate sustainable behaviour change. A focus on habits is crucial to designing effective DBCIs for three main reasons. Firstly, habitual behaviour is common in everyday life in multiple domains. Secondly, we have shown that reasoned-action theories and corresponding Type 2 techniques alone are unable to achieve lasting behaviour change in the face of strong habits. Thirdly, we have identified multiple opportunities for pervasive computing technology to deliver interventions that can target both Type 1 and Type 2 processes.

We note that a lack of theory can impact on intervention efficacy and generalisability. To illustrate how three sets of theories (Dual Process Theory, modern habit theory and Goal Setting Theory) suggest habits may operate, we integrated them into a descriptive framework, the Habit Alteration Model. We outlined a set of possible interventions based on the model, and generated a set of design principles to guide these interventions. The model, in contrast to most DBCI theory and research, demonstrates how both Type 1 nonconscious and Type 2 conscious processes can contribute to stable behaviour change.

The key technical challenges for DBCIs remain accurate and fast detection of context, and the familiar problem of evaluation. Other challenges remain in addressing outstanding theoretical issues, in robust testing of the suggested strategies, in avoiding reactance and in determining the most effective combination of the techniques for any given domain and implementation technology.

In conclusion, we strongly recommend that DBCI researchers consider the role of habits in the behaviour they are targeting, and consider the implications of the HAM to design interventions that target Type 1 processes alongside Type 2 processes. We look forward to a rich era of DBCI habitual behaviour change research.

ACKNOWLEDGMENTS

The authors would like to thank the reviewers for their constructive and insightful comments. Thank you to Robert J Hendley, Dr Rowanne Fleck and Diego Garaialde for their helpful comments. Special thanks to Dr Lindsay MacDonald Vermeulen for her help with the graphical representation of the HAM.

REFERENCES


Tegan Blackburne, Alexandra Rodriguez, and Stuart John Johnstone. 2016. A Serious Game to Increase Healthy Food Consumption in Overweight or Obese Adults: Randomized Controlled Trial. *JMIR Serious Games* 4, 2 (2016), e10. https://doi.org/10.2196/games.5708


Mark E Bouton. 2014. Why behavior change is difficult to sustain. *Preventive medicine* (June 2014). https://doi.org/10.1016/j.ypmed.2014.06.010


Thomas A. Brunner and Michael Siegrist. 2012. Reduced food intake after exposure to subtle weight-related cues. *Appetite* 58, 3 (June 2012), 1109–1112. https://doi.org/10.1016/j.appet.2012.03.010


Digital Behaviour Change Interventions to Break and Form Habits


Digital Behaviour Change Interventions to Break and Form Habits


Natalia S. Lawrence, Jamie O’Sullivan, David Parslow, Mahmood Javid, Rachel C. Adams, Christopher D. Chambers, Katarina Kos, and Frederick Verbruggen. 2015. Training response inhibition to food is associated with weight loss and reduced energy intake. *Appetite* 95 (Dec. 2015), 17–28.  https://doi.org/10.1016/j.appet.2015.06.009


Digital Behaviour Change Interventions to Break and Form Habits


Sheina Orbell and Bas Verplanken. 2015. The strength of habit. Health psychology review

Donald A. Norman and Tim Shallice. 1986. Attention to action


Jane C Walsh, Teresa Corbett, Michael Hogan, Jim Duggan, and Abra McNamara. 2016. An mHealth Intervention Using a Smartphone App to Increase Walking Behavior in Young Adults: A Pilot Study. *JMIR mHealth and uHealth* 4, 3 (Sept. 2016), e109. [https://doi.org/10.2196/mhealth.5227](https://doi.org/10.2196/mhealth.5227)


Brian Wansink. 2010. From mindless eating to mindlessly eating better. *Physiology & behavior* 100, 5 (July 2010), 454–63. [https://doi.org/10.1016/j.peh.2010.05.003](https://doi.org/10.1016/j.peh.2010.05.003)


Reinout W Wiers, Carolin Eberl, Mike Rinck, Eni S Becker, and Johannes Lindenmeyer. 2011. Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychological Science* 22, 4 (April 2011), 490–7. [https://doi.org/10.1177/0956797611400615](https://doi.org/10.1177/0956797611400615)


J M Williams, A Mathews, and C MacLeod. 1996. The emotional Stroop task and psychopathology. *Psychological Bulletin* 120, 1 (July 1996), 3–24. [https://doi.org/10.1037/0033-2909.120.1.3](https://doi.org/10.1037/0033-2909.120.1.3)


