

# Understanding long-term behaviour change techniques: a mixed methods study

Philip Cash <sup>a</sup>, Agata Wrobel <sup>b</sup>, Anja Maier <sup>b,c</sup> and John Paulin Hansen <sup>b</sup>

<sup>a</sup>Northumbria University, London, UK; <sup>b</sup>DTU Technical University of Denmark, Kongens Lyngby, Denmark; <sup>c</sup>University of Strathclyde, Glasgow, UK

## ABSTRACT

Long-term behaviour change is essential to many societal and personal challenges, ranging from maintaining sustainable lifestyles to adherence to medical treatment. However, prior research has generally focused on interventions dealing with bounded, present-tense, and discretely measurable behaviour change problems, evaluated via relatively short-term trials. This has led to a skewed prioritisation of behaviour change techniques and left a critical gap in design guidance. Hence, there is an urgent need to (i) examine how behaviour change techniques can be abstractly prioritised and (ii) related to contextual, embodied interventions during long-term behavioural design. We address this need using a Delphi survey method with 12 international experts on behavioural intervention complemented by a reanalysis of over 100 real-world cases. This provides the basis for examining how experts prioritise the Behaviour Change Technique Taxonomy (BCTT) for the long-term, as well as how this corresponds to real-world long-term interventions. Based on this we provide essential, and as a first, guidance for long-term behavioural design as well as contributing to wider research on how to deal with the demands of long-term behaviour change.

## ARTICLE HISTORY

Received 18 April 2023  
Accepted 18 June 2023

## KEYWORDS

Behavioural design; design for health; behaviour change techniques; long-term behaviour change; Delphi method

## Introduction

Long-term behaviour change is essential to addressing many societal challenges, including sustainable ways of living and healthier lifestyles (Green and Vergragt 2002; Kelly and Barker 2016), as well as more personal challenges, including medical treatment or diet (Jack et al. 2010; MacPherson et al. 2021). However, to ethically deliver and then maintain such long-term change requires behavioural designers to balance the dual demands of efficacy and effectiveness (Lilley and Wilson 2013; Mejía 2021; Michie, Atkins, and West 2015). Specifically, behavioural designers create interventions that are both contextually relevant and impactful as well as explicable and predictable, typically via the integration of design and behavioural science guidance (Niedderer, Clune, and Ludden 2017; Schmidt 2020). Yet, both design and behavioural science research focused on long-term behavioural change is limited (Kwasnicka et al. 2016; Pschetz and Bastian 2018; Schmidt and Stenger 2021).

**CONTACT** Philip Cash  philip.cash@northumbria.ac.uk  Northumbria University, School of Design Business and Law Building, 12 Falconar St, Newcastle upon Tyne NE2 1XA, UK

Behavioural design synthesises insights from design and behavioural science to deliver interventions that shape behaviour (Bucher 2020; Wendel 2013). Here, the abductive reasoning, creative reframing, and context sensitivity scaffolded by design (Dorst 2011; Kouprie and Visser 2009) are critical to identifying and delivering the most relevant and impactful interventions and then deploying and maintaining them in context. Complementary to this, the inductive/deductive reasoning, evidence based testing, and abstract theory supported by behavioural science (Dolan et al. 2014; West and Michie 2016) are essential to the detailed design, validation, and explanation of interventions with – at least somewhat – predictable outcomes. This creates a back-and-forth between design-led exploration and contextual embodiment and behavioural science-led explanation and validation. Hence, behavioural design is most effective when it integrates the best of design and behavioural science guidance (Cash et al. 2022b; Reid and Schmidt 2018).

In this context, designers repeatedly translate between abstract explanations of potential behaviour change techniques and their contextual embodiment in interventions, with projects faltering or failing when either side of this equation is neglected (Cash, Gram Hartlev, and Durazo 2017; Fogg 2009a). This creates a major challenge to long-term behavioural design due to two main deficits in current guidance. First, while research on abstract techniques is extensive their long-term validity is much less well understood. This is due to difficulties in conducting rigorous longitudinal studies and the diminishing impact of interventions over time (Hedin et al. 2019; Kwasnicka et al. 2016; Michie, Atkins, and West 2015), with limited and inconsistent evidence for the long-term effectiveness of many interventions (Gourlan et al. 2016; Williams and French 2011). Second, lists of techniques are often disconnected from the contextual realities faced by behavioural designers, being abstracted from the context of application as well as the embodiment of the interventions themselves (Hagger and Weed 2019; Schmidt and Stenger 2021). These deficits have the potential to derail long-term behavioural design efforts by simultaneously over-emphasising short-term techniques and de-emphasising the complexities of contextual intervention embodiment. Thus, there is an urgent need to (i) examine how behaviour change techniques can be abstractly prioritised and (ii) related to contextual, embodied interventions during long-term behavioural design.

Given this need, and limitations in current theory, we employ two complementary analyses to connect abstract prioritisation with contextual, embodied interventions used in practice. First, we examine how experts prioritise one of the most widely researched sets of behaviour change techniques, the Behaviour Change Technique Taxonomy (BCTT) (Michie et al. 2013). To this end we follow prior research on technique prioritisation when faced with limited evidence and theory (Kwasnicka et al. 2016; Samdal et al. 2017) by using the Delphi method (Vestjens et al. 2015; Willems et al. 2019). Second, we contextualise this by examining real-world embodied interventions in long-term behavioural design, based on the set of cases described by Cash et al. (2020). Together these explicitly connect design and behavioural science relevant considerations in this context.

## Theoretical background

Behavioural design integrates aspects of design and behavioural science to address behavioural and social challenges (Bucher 2020; Niedderer et al. 2016; Reid and Schmidt 2018). These range in scale from the personal to the societal and from the short- to the

very long-term (Maier and Cash 2022; Schmidt and Stenger 2021). This makes behavioural design a key means of delivering interventions that impact everything from maintaining sustainable food consumption to adherence to medical rehabilitation or increasing socially responsible behaviour (Hedin et al. 2019; MacPherson et al. 2021; Tromp, Hekkert, and Verbeek 2011). Thus, behavioural designers develop interventions, often comprising multiple design artefacts, which ‘explicitly and ethically realise positive behaviour, desired by both the individual and society’ based on behavioural theory (Khadilkar and Cash 2020, 521).

The behavioural design process typically involves five main phases that both individual and as a whole can serve to scaffold a ‘best of both’ balance between design and behavioural science guidance. For example, it is possible to see this interplay across the five phases described by Cash et al. (2022a): (i) **Identification and Framing**: emphasises framing and reframing of the problem/solution in conjunction with characterising the problem and developing empathetic understanding; (ii) **Mapping and Description**: balances the development of empathy and deep contextual understanding with the detailing of root causes and their explanation via behavioural theory; (iii) **Framing and Development of Interventions**: emphasises framing and reframing of the solution space to identify interventions that are both contextually relevant and impactful as well as valid and explainable; (iv) **Iterative Testing and Refinement**: balances creative design and refinement with validation focused behavioural evaluation; (v) **Scaling up, Launching and Maintenance**: emphasises context specific delivery, monitoring, and maintenance of interventions to realise the desired behaviour change.

During this process – and particularly Phase 3 – behavioural designers translate between abstract explanations of potential behaviour change techniques and their contextual embodiment in tangible interventions. This iterative translation balances abduction and reframing in search of the most contextually relevant and valuable intervention with the need to explain and subsequently predict and test the validity of interventions that are often deployed into critical contexts such as health or safety. However, to support this work designers must connect abstract explanations, typically associated with isolated behavioural change techniques, and contextual understanding of potential embodiments that often combine techniques (Bohlen et al. 2020; Schmidt and Stenger 2021). Hence, there is a need to understand both the abstract and embodied prioritisation of techniques relevant to the context of the design work. As such, the deficits highlighted in the introduction pose a major challenge to long-term behavioural design.

### **Long-term behavioural design**

Two major literature streams have developed that are relevant to supporting long-term behavioural design: processes for organising development (Cash, Gram Hartlev, and Durazo 2017; Fogg 2009a; Wendel 2013), and guidelines or techniques for translating behavioural theory into artefacts and interventions (Bhamra, Lilley, and Tang 2011; Lockton, Harrison, and Stanton 2010; Tromp and Hekkert 2014). Critically, while numerous aspects relevant to long-term behavioural design are highlighted across this literature (e.g. relating to habits or personal identity (Lockton, Harrison, and Stanton 2008; Michie et al. 2013)), there is a general focus on achieving intervention delivery and measurable impact within the context of a project, and hence, typically in the short-term (Khadilkar and Cash 2020; Schmidt and

Stenger 2021). Thus, processes and techniques for delivering long-term behavioural design are less well developed.

This creates an important gap in behavioural design knowledge, as insights from short-term development and measurement can not necessarily be directly translated to the long-term (Fogg 2009b; Schmidt and Stenger 2021). This is because long-term interventions are impacted by dynamic changes in, for example, the physical and social environment as well as personal behaviours and habits (Schmidt 2022; Schmidt and Stenger 2021). In terms of processes, this has led authors such as Schmidt and Stenger (2021) to highlight the need for the creation of more resilient solutions by embracing more diverse forms of evidence and applied foresight coupled with the iterative development of interventions within ecosystems over time. Similarly, Maier and Cash (2022) argue for an integrative view of interventions within a system, explicitly adopting a reflexive, agile approach able to monitor, reflect on, and react to a changing system over time. While these works feature potential ways forward in terms of long-term behavioural design processes, it is less clear how current techniques should be contextually understood and prioritised in guidance for long-term applications. This is despite authors such as Fogg (2009b) highlighting the need for general differentiation between the short- and long-term. Thus, there is a specific need to examine how techniques should be understood in the context of long-term behavioural design.

### ***Behaviour change techniques for long-term behavioural design***

Behavioural designers have access to a wide array of behaviour change techniques described in numerous lists, taxonomies, and frameworks (for overview see Maier and Cash (2022)). These range in scope and abstraction i.e. the degree to which they concretise the link between general behavioural theory and specific intervention design. For example, works such as the Behaviour Change Technique Taxonomy (BCTT) (Michie et al. 2013) or Mindspace (Dolan et al. 2014), are both broad and abstract, while the frameworks provided by Cash et al. (2020), Tromp, Hekkert, and Verbeek (2011), Bhamra, Lilley, and Tang (2011), or Lockton, Harrison, and Stanton (2010), offer various degrees of specificity in the link between abstract theory and contextualised intervention. While these works provide overall guidance to designers, they provide little support for the essential translation between abstract technique and contextually embodied intervention. In particular, few works explicitly address interactions between possible techniques and contextual factors that might impact prioritisation when designing and typically offer only limited guidance on technique embodiment in context. Further, none of these works offer in depth guidance on how their lists of behaviour change techniques should be prioritised for long-term interventions. Therefore, we build on the BCTT (the most widely recognised, researched, and applied set of abstract behaviour change techniques) as a starting point for our research. This list is also at the heart of many behavioural science-led projects, yet at the same time, is one of the most abstract. As such, contextualisation of this list offers great potential in developing the connection between design and behavioural science guidance.

The BCTT contains 16 clusters, with 93 individual BCTs, presenting a huge, yet abstract, scope for intervention design (Michie et al. 2013). This breadth has supported application across a wide range of long-term oriented settings, including rehabilitation (Meng et al. 2014), lifestyle change (Willems et al. 2019), elderly care (Vestjens et al. 2015; Walsh et al. 2021), and smoking cessation (O'Neill et al. 2018; Ubhi et al. 2016). However, much

of the research on BCTT application reflects trials adopting a primarily behavioural science approach and typically not addressing more design related contextual and embodiment related considerations. This has led to a situation where despite strong support for the validity and efficacy of the BCTs themselves there is still a high failure rate in real-world interventions (Hagger and Weed 2019; Schmidt and Stenger 2021). Further, Nieuwenhuisen et al. (2006) specifically highlight how personal context and environmental factors can act as facilitators or barriers to change, and it is generally acknowledged in the design literature that embodiment considerations can make or break intervention outcomes. Thus, in order to scaffold the translation between abstract and contextual guidance at the heart of behavioural design – and essential to integrating the best of both design and behavioural science – there is a need to better understand BCTT prioritisation in relation to intervention embodiment and change over time (following recent work on the need to consider temporality in design (Pschetz and Bastian 2018)), as well as interaction with changes in the physical, social, and technological environment (following recent work highlighting interactions between intervention and context (Cash et al. 2020; Schmidt and Stenger 2021)).

### **Research framework**

Given the needs outlined in the previous section, our research framework starts by building on the 16 major clusters in the BCTT (Michie et al. 2013). We subsequently examine how these BCTT clusters can be abstractly prioritised and related to contextual, embodied interventions during long-term behavioural design. Specifically, we compare and contrast abstract and embodied prioritisation and implementation of the BCTT clusters with respect to long-term interventions, susceptibility to change over time, and interaction with changes in the physical, social, and technological environment. Here, long-term is defined as lasting three months or longer, following research on general habit formation (Lally et al. 2010) as well as more specific studies in, for example, health related contexts such as stroke rehabilitation (Teasell et al. 2012). This leads to two main research questions that explicitly connect design and behavioural science relevant considerations for long-term behavioural design:

**RQ1:** What are the most important BCTT clusters for long-term behavioural design?

**RQ2:** How susceptible are these BCTT clusters to change over time and in different contexts?

### **Methodology**

Given the nature of the research need outlined in the previous section our methodology builds out from a classic Delphi study, which we first elaborate by adding additional context related questions, and then complement via a systematic analysis of real-world cases. This mixed-methods methodology seeks to connect abstract and contextual insights and thus draws together approaches found in design i.e. case-based ranking, and behavioural science i.e. expert-based ranking.

First, the multi-round Delphi method has proven a reliable means of eliciting knowledge from a varied group of experts (as found in behavioural design) (Okoli and Pawlowski 2004). This has four features: (1) anonymity via using questionnaires, (2) iteration to achieve consensus, (3) controlled feedback, and (4) statistical aggregation to distil the final response

(Rowe, Wright, and Bolger 1991). The Delphi method is especially appropriate when there are limitations in available theory or empirical data – as is the case in long-term behavioural design – with experts offering insights not accessible via other means, based on years of collective experience. Here, experience from diverse experts is brought together to provide a distillation of best practice that can help guide others. This approach has been widely used in the context of BCTT ranking (Vestjens et al. 2015) and seen more limited application in the design domain when seeking to elicit insights from across expert groups (Adams and Meyer 2009; McMahon and Bhamra 2015). Further, the type of ranking produced by the Delphi method is readily comparable to other rankings or similarly structured guidance, which are already common in the context of behavioural design (Bhamra, Lilley, and Tang 2011; Lockton, Harrison, and Stanton 2010; Tromp, Hekkert, and Verbeek 2011).

Second, providing a contextual, embodied counterpoint to the abstract Delphi results, we systematically examined real-world interventions targeting long-term change. This built on the set of cases previously described by Cash et al. (2020). This type of approach has been widely used in the design literature (Kelders et al. 2012; Lockton, Harrison, and Stanton 2010; Tromp, Hekkert, and Verbeek 2011), and allows for the examination of how BCTT clusters are embodied in real-world interventions. This provides critical nuance to the Delphi method results. For example, while BCTT clusters are typically ranked individually they are frequently embodied in combination in actual interventions and hence understanding their real-world implementation is key to scaffolding translation between ranking and reality. This is important as it emphasises the guiding and explanatory role of such rankings as well as that they should not form a prescriptive straitjacket. Rather there is significant scope and need for design exploration and reframing in the translation between abstract ranking and contextual embodiment.

This combination of methods allows us to distil insights relevant to designers working with long-term change by connecting design and behavioural science perspectives, whilst minimising the ethical complexities associated with behavioural trials and long-term data collection. To provide a common basis for building this connection we take the BCTT clusters as a widely understood and applied starting point (Michie et al. 2013; Vestjens et al. 2015; Willems et al. 2019). The 16 BCTT clusters are: (1) goals and planning, (2) feedback and monitoring, (3) social support, (4) shaping knowledge, (5) natural consequences, (6) comparison of behaviour, (7) associations, (8) repetition and substitution, (9) comparison of outcomes, (10) reward and threat, (11) regulation, (12) antecedents, (13) identity, (14) scheduled consequences, (15) self-belief, and (16) covert learning.

## **Sampling**

### **Delphi panel**

Following the considerations outlined by Cash et al. (2022a) we first bounded the scope of this work by sampling experts from across the behavioural design spectrum, from designed to behavioural science-led (Reid and Schmidt 2018). This provides the basis for deriving generalised insights by seeking consensus across a diverse sample, and a specific set of criteria for a purposive sampling schema – aligned with the Delphi method (Onwuegbuzie and Leech 2007; Sumsion 1998). Specifically, we identified 12 international experts on behavioural intervention, based on authors' networks, following prior BCT focused Delphi studies (Garvin and Simon 2017; Vestjens et al. 2015). Here our sampling criteria was

**Table 1.** Overview Delphi panel.

Sample	Countries	Age	Highest degree	Employment	Current focus
Design-led	Denmark (5) Portugal (1)	Mean 32 SD 7	PhD (2) Masters (4)	Private (5) University (1)	Founder of behavioural design company; project manager; physical/digital designer; UX; wayfinding
Behavioural science-led	Denmark (3) Germany (2) Portugal (1)	Mean 40 SD 9	PhD (2) Masters (3) Doctor (1)	Government (2) Private (3) University (1)	Neuropsychological assessment; rehabilitation; neurorehabilitation; child psychology

expertise in long-term behavioural design, development of interventions in this context, and familiarity with the BCTT i.e. they were able to distinguish the BCTT clusters in principle, distinct from embodied interventions. Experts were drawn from a range of backgrounds including behavioural design, behavioural science, and healthcare, all working with long-term behaviour change. For the purposes of this research these were broadly categorised as predominantly design- ( $n = 6$ ) or behavioural science-led ( $n = 6$ ) (Reid and Schmidt 2018). Amongst the six design-led experts two had PhDs and all worked in practice ranging from founder of an internationally recognised behavioural design consultancy known for drawing on both design and behavioural science guidance to experienced UX, wayfinding, and digital/physical product design. Amongst the six behavioural science-led experts two had PhDs and one was a medical doctor and all worked in practice ranging from head of a neurological and research department to stroke and other rehabilitation. All experts had at least a master's degree as well as multiple years of experience in practice. The panel is summarised in Table 1.

Based on this sampling approach, 12 experts completed the first round and 9 (75% – meeting the 70% threshold for acceptability (Sumsion 1998)) the second round. While Delphi studies can vary in sample size there is a need to ensure balanced representation of groups within the sample as well as sufficient expertise in use of the BCTT. As such, our identified sample of 12 experts, while small, is aligned with many other recent studies specifically dealing with BCT prioritisation (Davies, Martin, and Foxcroft 2016; McCarthy et al. 2020; O'Neill et al. 2018; Vestjens et al. 2015; Walsh et al. 2021; Willems et al. 2019), as well as similar Delphi type studies in the design domain. For example, McMahon and Bhamra (2015) had a sample of 19 experts when examining competencies for social sustainability and Adams and Meyer (2009) complemented a sample of eight experts with a wider secondary sample. Across contexts, there is an acknowledgement that relevant experts are difficult to access and that valid Delphi results can be derived from small sample sizes. Specifically, the heterogeneous nature of the sample allows insights to be derived from small groups and hence is in line with guidelines in this context (Clayton 1997; Humphrey-Murto et al. 2020).

### *Real-world cases*

To provide a contextual, embodied counterpoint to the abstract Delphi results, and hence scaffold designer's translation between design and behavioural science guidance we systematically examined real-world interventions stemming from long-term behavioural design. To do this we re-analysed the set of 139 cases previously identified by Cash et al. (2020).

This set of cases was derived from across domains (including regulation, safety, and security; health habits, activity, and food; personal routines and finances; sustainability; and volunteering (Cash et al. 2020, fig. 2a)) and included numerous short- and long-term interventions, drawing on reports published by specialist behavioural design companies, internal reporting from companies, intervention databases, and interventions described in academic literature. Key inclusion criteria for this sample were data quality (i.e. describing both problem and solution), solution quality (i.e. inclusion of behavioural success measures), validity (i.e. the testing methodology was robust), and ethics (i.e. following ethical best practice).

While Cash et al. identified 218 interventions based on their coding schema some of these combined BCTT clusters and hence we identified 321 unique BCT interventions across the dataset. For example, some cases comprised multiple artefacts and/or service components (e.g. an app, a messaging service, and a discussion group all packaged together) each drawing on a distinct BCTT cluster in their delivery. This allowed us to contrast the ranking of individual BCTT clusters in the Delphi method with both their individual application in real-world interventions and their combination in this context. Each intervention was coded in terms of both its focus (dealing with the short- or long-term) and the BCTT cluster it utilised. Of the 321 interventions 123 focused on short-term change while 198 focused on long-term. Here, short-term was defined as interventions focused on a one-time event or rare behaviour. For example, choosing an insurance policy or deciding whether to be an organ donor were considered rare short-term decisions. In contrast, long-term was defined as interventions focused on repeated or frequent behaviours (e.g. daily or weekly). For example, rehabilitation exercises after surgery or a stroke must be done most days over an extended period and hence, they are long-term. This also supported the generalisation of the results by contrasting the abstract Delphi ranking with interventions drawn from across different long-term behavioural design contexts. Intercoder reliability was evaluated in two rounds by the first and second author. In round one 10% of interventions were analysed and revealed 83% agreement – primarily due to ambiguity in the coding of the Associations BCTT cluster. This ambiguity was clarified and round two evaluated a further 10% to achieve 98% agreement. Based on the high level of agreement at this stage the rest of the cases were coded by the second author. Ranking results were then calculated based on the long-term uses of a BCTT cluster normalised against the total uses of that BCTT cluster to produce a score from 1-5. This allowed for comparison with the Likert based rankings.

Overall, this comparison with the Likert based rankings serves to contextualise the expert data because while the expert questions could focus on abstract validity the case data reflects real world usage, which can be impacted by ease of delivery or maintenance over time, preferred delivery modes, or other pragmatic design factors. As such, the two datasets complement one another by providing abstract and contextual counterpoints in validity and usage.

### **Delphi procedure**

Qualtrics software was used for the first round of the Delphi study, which took the form of an online questionnaire. Following informed consent and demographic information, participants were asked to review each of the 16 BCTT clusters based on a summary overview and image, as illustrated in Figure 1. They then answered three main questions for each cluster



#### 4. Shaping Knowledge: Summary



In this section, you will be asked questions about the category of Behaviour Change Techniques (later referred to as BCTs) called **Shaping Knowledge**.

**Shaping Knowledge** category comprises BCTs that revolve around providing information on the antecedents of the behaviour (including suggesting alternative explanations) and instructions on how to perform the behaviour.

The full list includes the following BCTs:

- Instruction on how to perform a behaviour
- Information about antecedents (e.g. social and environmental situations, emotions, cognitions, that reliably predict performance of the behaviour)
- Re-attribution
- Behavioural experiments (advise on how to identify and test hypotheses about the behaviour, its causes and consequences)

**Example(s):**

**Re-attribution** means to elicit perceived causes of behaviour and suggest alternative explanations. For example, if the person attributes their over-eating to the frequent presence of delicious food, suggest that the 'real' cause may be the person's inattention to bodily signals of hunger and satiety.

**Figure 1.** Example of BCTT cluster presentation for cluster: 4) shaping knowledge.

using a five-point Likert scale (rated from 'not at all important' to 'extremely important' for Question 1 and from 'not at all' to 'A great deal' for Questions 2 and 3):

- (1) In your opinion, how important is [BCTT cluster] for long-term behaviour change?
- (2) In your opinion, how much does the effectiveness of [BCTT cluster] change over time? (For example, would it work equally well in a course of a month, 3 months, 6 months, a year?)
- (3) In your opinion, how susceptible is [BCTT cluster] to changes in each of these factors? (In other words, if person's physical, social, or technological environment changes, would it affect this behaviour change technique?)
  - (i) Physical environment (e.g. home, hospital, surroundings)

- (ii) Social environment (e.g. family, friends, social network)
- (iii) Technological environment (e.g. available equipment, such as computer, tablet, phone, or rehabilitation aids)

Based on the results from the first round, a second questionnaire was developed specific to each participant and delivered via email. Here, each participant received a spreadsheet overview of their own answers in comparison to the round 1 median and were then invited to reconsider their answer and offer comments. Following standard Delphi method practice (Willems et al. 2019), this provided feedback on the responses from the first round, as a basis for participants to reflect on and potentially alter their previous answers. Thus, participants had the chance to either change their prior answer or to give additional explanation of their reasoning. The same explanation of each BCTT cluster was provided as in the first round.

### Data analysis

Based on the survey results the level of consensus was calculated for each question. This first evaluated the whole sample, and then also the design and behavioural science sub-samples. Following prior works in this context, consensus was evaluated as  $\geq 50\%$  agreement around a median (Gracht 2012). This allowed for a more nuanced understanding of the results in relation to RQ2 and provided a basis for ranking the BCTT clusters with respect to importance for long-term behavioural design in relation to RQ1.

The percentage consensus and median are summarised for all BCTT clusters in Table 2 (as an illustration of the analysis of survey Question 1) and in Appendix Tables 7, 8, and 9 for Questions 2 and 3. For Question 1 good consensus was found across the sample for all but three BCTT clusters. For Questions 2 and 3 consensus was found for the majority of BCTT clusters but also showed divergence between the design and behavioural science sub-samples (discussed in more detail in the results). Once consensus was evaluated, rankings were calculated based on the mean response for each BCTT cluster, following prior BCT rankings (Vestjens et al. 2015).

**Table 2.** Overview of consensus results for ranking of importance for long-term behavioural design.

BCT	Sample					
	Combined		Design		Behavioural science	
	Consensus	Median	Consensus	Median	Consensus	Median
1. Goals and planning	67	4.0	50	4.5	83	4.0
2. Feedback and monitoring	50	4.5	50	4.5	50	4.5
3. Social support	58	4.0	83	4.0	33	4.0
4. Shaping knowledge	50	4.0	67	4.0	50	3.5
5. Natural consequences	50	3.0	50	3.0	50	3.0
6. Comparison of behaviour	42	2.5	50	3.5	50	2.0
7. Associations	50	4.0	50	3.5	50	4.0
8. Repetition and substitution	58	4.0	67	4.0	50	4.5
9. Comparison of outcomes	50	3.0	50	3.5	67	3.0
10. Reward and threat	67	3.0	67	3.0	67	3.0
11. Regulation	50	3.5	83	4.0	67	3.0
12. Antecedents	50	4.0	50	3.5	50	4.0
13. Identity	42	4.0	50	4.0	67	3.0
14. Scheduled consequences	42	3.5	50	2.5	50	4.0
15. Self-belief	58	4.0	50	4.5	67	4.0
16. Covert learning	67	3.0	83	3.0	50	3.0

**Table 3.** Ranking of importance for long-term behavioural design.

BCT	Sample			
	Combined	Design	Behavioural sci.	Cases
	Rank	Rank	Rank	Rank
<i>15. Self-belief</i>	1	2	1	1
<i>8. Repetition and substitution</i>	2	1	7	1
<i>2. Feedback and monitoring</i>	3	2	4	4
<i>1. Goals and planning</i>	4	5	1	5
3. Social support	5	4	4	7
4. Shaping knowledge	6	7	3	8
11. Regulation	7	11	8	1
13. Identity	8	4	10	14
7. Associations	9	14	5	11
9. Comparison of outcomes	10	9	11	9
6. Comparison of behaviour	11	16	9	6
<i>10. Reward and threat</i>	<i>12</i>	<i>11</i>	<i>9</i>	<i>10</i>
<i>14. Scheduled consequences</i>	<i>13</i>	<i>9</i>	<i>16</i>	<i>12</i>
<i>12. Antecedents</i>	<i>14</i>	<i>7</i>	<i>9</i>	<i>15</i>
<i>5. Natural consequences</i>	<i>15</i>	<i>11</i>	<i>14</i>	<i>13</i>
<i>16. Covert learning</i>	<i>16</i>	<i>15</i>	<i>9</i>	<i>16</i>

## Results

In answer to our research questions, we first rank the importance of BCTT clusters for long-term behavioural design, before examining the varied expert views on the susceptibility of these clusters to change over time and in different contexts.

### Ranking importance for long-term behavioural design

In answer to RQ1, consensus was reached for 13 out of 16 BCTT clusters. The importance of all clusters for long-term behaviour design is summarised in Table 3, with the most and least relevant denoted in italics. The most relevant BCTT clusters-based on the combined and case means-were self-belief (mean = 4.61), repetition and substitution (mean = 4.50), feedback and monitoring (mean = 4.35), and goals and planning (mean = 4.27); while the least relevant were scheduled consequences (mean = 3.11), antecedents (mean = 3.05), natural consequences (mean = 2.94), and covert learning (mean = 2.33). The full details of the mean scores and ranks for all sub-samples is provided in Appendix Table 10 (here columns are organised sub-samples associated with the six design-led experts, six behavioural science-led experts, and the real-world cases; mean results were used for consistent comparability across sub-samples), while Table 11 provides the supporting details of the case mean calculation.

### Contrasting factors influencing long-term interventions

The answer to RQ2 was less clear because, while sub-sample (design and behavioural science) consensus was relatively high (circa 13 out of 16 BCTT clusters with  $\geq 50\%$  agreement for all questions), agreement between the samples diverged as highlighted in comments by the experts during the second round of the Delphi method (see Appendix Tables 7–9). Results are summarised in Tables 4a and 4b.

**Table 4a.** Ranking of susceptibility to other factors (Questions 2–3).

BCT	Time		BCT	Physical environment	
	D Rank	B Rank		D Rank	B Rank
10. Reward and threat	1	2	1. Goals and planning	1	6
1. Goals and planning	2	4	9. Comparison of outcomes	1	6
2. Feedback and monitoring	3	8	7. Associations	3	2
5. Natural consequences	4	14	11. Regulation	4	4
3. Social support	5	4	2. Feedback and monitoring	4	9
4. Shaping knowledge	5	4	3. Social support	4	10
11. Regulation	5	4	15. Self-belief	4	10
16. Covert learning	5	12	4. Shaping knowledge	8	10
8. Repetition and substitution	9	1	12. Antecedents	9	1
7. Associations	9	14	8. Repetition and substitution	9	3
13. Identity	11	8	16. Covert learning	11	8
6. Comparison of behaviour	12	8	14. Scheduled consequences	11	10
12. Antecedents	13	3	5. Natural consequences	11	15
15. Self-belief	13	8	10. Reward and threat	14	4
9. Comparison of outcomes	13	13	6. Comparison of behaviour	15	14
14. Scheduled consequences	13	14	13. Identity	15	16

Sample denoted by Design (D) and Behavioural Science (B).

**Table 4b.** Continued.

BCT	Social environment		BCT	Technical environment	
	D Rank	B Rank		D Rank	B Rank
3. Social support	1	1	2. Feedback and monitoring	1	1
7. Associations	2	14	1. Goals and planning	1	3
6. Comparison of behaviour	3	3	4. Shaping knowledge	3	3
4. Shaping knowledge	3	6	7. Associations	3	5
1. Goals and planning	5	2	9. Comparison of outcomes	3	14
2. Feedback and monitoring	5	8	16. Covert learning	6	7
9. Comparison of outcomes	5	8	14. Scheduled consequences	6	9
10. Reward and threat	5	10	8. Repetition and substitution	8	5
13. Identity	9	3	6. Comparison of behaviour	8	9
15. Self-belief	9	3	12. Antecedents	10	2
16. Covert learning	9	7	10. Reward and threat	10	9
8. Repetition and substitution	9	10	15. Self-belief	12	14
5. Natural consequences	9	16	13. Identity	13	8
12. Antecedents	14	12	3. Social support	13	9
14. Scheduled consequences	15	12	5. Natural consequences	13	16
11. Regulation	16	14	11. Regulation	16	9

Sample denoted by Design (D) and Behavioural Science (B).

In terms of susceptibility to change over time, there was a greater than eight rank disagreement on natural consequences (rank = 4 design v. 14 behavioural science), covert learning (rank = 5 design v. 12 behavioural science), and antecedents (rank = 13 design v. 3 behavioural science). Similarly, in terms of susceptibility to change in physical environment there was a greater than eight rank disagreement on antecedents (rank = 9 design v. 1 behavioural science) and reward and threat (rank = 14 design v. 4 behavioural science), social environment on associations (rank = 2 design v. 14 behavioural science), and technological environment on comparison of outcomes (rank = 3 design v. 14 behavioural science) and antecedents (rank = 10 design v. 2 behavioural science). Full detail of all means is provided in Appendix Tables 12 and 13.

### **Linking abstract and contextual results**

To elaborate insights relevant to designers working with long-term behavioural design we also explored the experts' rationale for their rankings and further connected these abstract rankings with deeper analysis of the real-world cases. Here, justifications for differences in these ranks focused on experience and differing foci on technologies or specific types of people being influenced. For example, Expert 8 highlighted how in their experience: 'Your social environment has a great deal to say about beliefs of natural consequences and how one should interpret social rules and/or real health consequences'. Similarly, Expert 2 emphasised how 'In my experience, associative techniques and cueing can be difficult to transfer from one setting to another. Often, training needs to start over when transitioning from one setting to another (e.g. rehabilitation centre to home)' and further that 'I believe that planning and goal setting are social actions to an extended degree. Relatives, friends, etc. are very important for setting goals and following up on them'. This focus on contextual understanding, and lack of clear guidance on the influence of context factors, thus has the potential to explain differing opinions between sub-samples. However, all experts agreed that BCTT clusters were significantly impacted by these factors and that they were essential to consider when designing long-term interventions.

Unpacking these results further revealed several contextual design considerations around the BCTT clusters identified as most important in Table 3. Here multiple experts highlighted how embodiment of BCTT clusters in real-world interventions needed to change over time to maintain relevance and effect. For example, Expert 8 noted that feedback based interventions needed to adjust over time as behaviour become habitual: 'after some time, feedback stops being the initiator of new behaviour, the point where it's a habit' and further that as people move from one context to another over time the role of an intervention can change: 'social support is often based on where you are and getting out of a hospital after days of care will greatly change your interaction with it'. Expert 12 also explained that in their view

the effect of planning and goal-setting can become less effective over time ... depending on a number of factors such as; Are the goals achievable? Are they realistic? Does the person experience some direct benefit of sticking to the plan (or some direct negative consequence if not sticking to it)?

This need to account for development over time was also highlighted by the behavioural science-led experts. For example, Expert 1 stated that 'a long period of rehab goal and planning are very effective for the patients. As long as the patient evaluates and continuously changes the goals during the rehab'. As such, there is a critical need for design exploration in the translation from abstract BCTT cluster into interventions not only in the immediate context but with consideration to change over time.

In addition to temporal considerations, several design-led experts highlighted embodiment and contextual factors that impact how BCTT clusters might be incorporated into real-world interventions. For example, Expert 7 pointed out this contextual dependency in embodiment: 'goals and planning are very technology correlated, for example, by digital calendars, reminders etc'. This was also emphasised by Expert 8: 'setting out a plan/goals could, for example, be based on your fridge or shoes by the door, which makes these elements greatly influencing by the physical surroundings' and that 'the only thing that allows you to compare yourself is the social environment'. These considerations were also

mirrored by the behavioural science-led experts. For example, Expert 5 explained how ‘feedback through technological devices will have a greater impact than a therapist’s opinion, patients believe more in objective measures ... [but] ... if the technology changes, patients will not be able to compare the outcomes with each other’. However, they also highlighted how many BCTT clusters can be implemented through varied means. For example, Expert 2 stated that ‘technological solutions may assist identity techniques, but I do not think that the effectiveness of these techniques relies on technological solutions’. Overall, these results emphasise the importance of the translation from abstract BCTT cluster to embodied interventions that are meaningful and effective in context and points to the need for a deep understanding of both the design and behavioural science considerations that might influence these outcomes.

Connecting these insights to more detailed analysis of the real-world cases, two additional considerations are brought to the fore. First, most cases employed multiple interventions. Here, the total number of interventions that were distinct in terms of both BCTT cluster and their design embodiment was three across cases, while the average number when only distinguishing by BCTT was two. This highlights the high degree of variation and design scope when translating from abstract cluster to embodied intervention. Second, as evidenced by the difference in averages noted above many cases chose to embody a single BCTT cluster in multiple distinct interventions, combined clusters via multiple interventions, combined clusters in a single intervention, or some combination of these. The complexity of this translation from general guidance – often associated with one or two main BCTT clusters – and contextual embodiment is illustrated in Figure 2, which shows how many cases employed two or more interventions. This can be further understood by examining examples of interventions associated with each BCTT cluster, as summarised in Table 5. This gives an overview of the diversity of possible embodiments as well as the potential for combinations and synergies across BCTT clusters through smart design (Note: no long-term cases used the covert learning cluster so this is not included in Table 5).

## Discussion

Our research provides two main contributions addressing the need outlined in the introduction to (i) examine how behaviour change techniques can be abstractly prioritised (linked to RQ1) and (ii) related to contextual, embodied interventions during long-term behavioural design (linked to RQ2). These are primarily built on the results distilled in Tables 5 and 6. Table 6 brings together the abstract findings while Table 5 illustrates the diversity of contextual embodiments related to these.

The first main contribution is based on answering RQ1 (what are the most important BCTT clusters for long-term behavioural design?) by providing an overall, abstract ranking of BCTT clusters in Table 6. This addresses an important gap in research on behaviour change techniques in general (Kwasnicka et al. 2016), as well as highlighting how long-term BCTT prioritisation differs from more short-term focused rankings. For example, examining the real-world case data from Cash et al. (2020) for short-term usage there was a focus on antecedents and associations, which significantly contrasts with their low ranking in Table 6. Similarly, Watson et al.’s (2021) work on handwashing interventions also highlighted antecedents as well as comparison of behaviour and natural consequences, which are all in the lower half of the ranking in Table 6. Evaluating the credibility of our findings

**Table 5.** Example interventions associated with each BCTT cluster drawn from the cases.

BCT	Example 1	Example 2	Example 3
15. Self-belief	Embodied in encouraging text messages saying that 'you can succeed' and that 'you belong' in the intervention context.	Embodied in motivational text messages containing a variety of different messages and including e.g. public testimonials.	NA
8. Repetition and substitution	Embodied in a videogame where the user moves a character by moving themselves with an accelerometer to help them practice an activity.	Embodied in a guide helping people to select sustainable food and hence support more general consideration of sustainable food products.	Embodied in a tooth brushing game where children practice with guidance from changing tones and visual images that provide a sense of control.
2. Feedback and monitoring	Embodied in a computer game where, fish size displayed the progress toward a target while fish facial expression showed if the goal was reached.	Embodied in a home touch screen display, which provides ambient feedback on electricity usage and associated cost via colour-coding.	Embodied in a shoe accessory with a light that brightens the more the wearer walks and slowly dims when the wearer remains stationary.
1. Goals and planning	Embodied in a saving plan consisting of (i) an approach to increase contribution, (ii) savings start after an increase in salary, and (iii) saving percent increase at each scheduled raise.	Embodied in a safe tool with suggested activities (e.g. play football) for the weekend for youth, requiring them engage with the tool, and then encouraging them to contact friends to do it with.	Embodied in a process where people write down three plans for when, how and where they are going to exercise and three barriers that might challenge this and how to cope with them.
3. Social support	Embodied in a competition about energy use across floors in a building and making one person responsible for advocating energy saving for a day.	Embodied in an online consultation providing a virtual social training session that is complemented by other activities related to the user's context.	Embodied in an online social 'sharing group' where users can comment step counts and get support from others.
4. Shaping knowledge	Embodied in a set guidance promoted via a media campaign using social media, blogs, PR, governmental webpages, and a specific app.	Embodied in weekly facilitated group meetings complemented by information leaflets on e.g. the benefits of exercising.	Embodied in a physical visit by a student who provided information about the current initiative to reduce water consumption.
11. Regulation	Embodied in a virtual garden representing physical activity with flowers as a metaphor and with a pleasing aesthetic so users won't change it.	NA	NA
13. Identity	Embodied in a personal group where users can send their step count and comments to any/all of their fitness buddies and see buddies' progress with respect to goals, amongst other things.	Embodied in a virtual platform where users can see a changing self-image of their development from smoker to non-smoker.	Embodied in a virtual garden representing physical activity with flowers as a metaphor and opportunities for users to identify with and share information about the garden with others.

(continued).

**Table 5.** Continued.

BCT	Example 1	Example 2	Example 3
7. Associations	Embodied in a game for elementary school children where users have to make recipes and eat more fruit, juice and vegetables to win.	Embodied in messages tailored to user's personal profile and connecting desired outcomes with specific prompts.	Embodied in a mobile app that reminds office workers to take break from the screen via messages linked to getting up from the desk.
9. Comparison of outcomes	Embodied in a programme where teenagers mentor toddlers through a 20-week course to see the outcome of early, unplanned pregnancy.	Embodied in a persuasive message where student actors and a doctor (also an actor) talk about a procedure and misbeliefs regarding it.	Embodied in an explanation showing users exactly what they are donating for and the specific difference in outcomes their donation makes.
6. Comparison of behaviour	Embodied in an ambient display worn on the wrist that notifies users in near real-time when other people in their social group are physically active.	Embodied in a virtual competition showing how much your colleagues has been drinking (water), visually represented by a growing tree.	Embodied in a Facebook application designed to provide social and competitive context for daily pedometer readings.
10. Reward and threat	Embodied in receipt-based lottery for individuals and a collective prize for the city if the amount of money spent on card exceeds a certain amount.	Embodied in rewarding messages highlighting where calorie intake was lower than the previous day.	NA
14. Scheduled consequences	Embodied in a pre-planned congratulatory message delivered on day 1 of quitting smoking.	NA	NA
12. Antecedents	Embodied in a redesigned prescription chart, which should be confirmed every third day and has an integrated checklist. 2.	Embodied in a 'smart' shopping cart with a LED light bar displaying various information on usage and comparison with social norms.	Embodied in sound and tactile cues and reminder built into a desk to trigger users to adjust it to their height.
5. Natural consequences	Embodied in a displaying converting the number of printed pages to equal number of trees or CO2.	Embodied in a documentary with the potential consequences of texting and driving.	NA



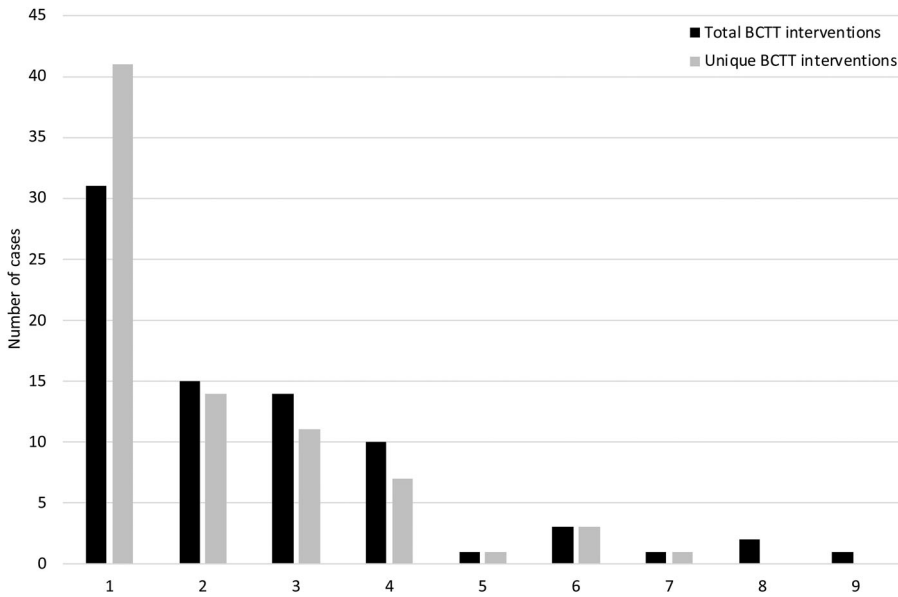


Figure 2. Overview of number of interventions used in each case.

Table 6. Ranking of importance versus overall susceptibility to change over time and contextual factors.

BCT	Overall importance		Overall susceptibility	
	Rank	Mean	Rank	Mean
<b>15. Self-belief</b>	<b>1</b>	<b>4.61</b>	<b>12</b>	<b>2.9</b>
<b>8. Repetition and substitution</b>	<b>2</b>	<b>4.50</b>	<b>5</b>	<b>3.3</b>
<b>2. Feedback and monitoring</b>	<b>3</b>	<b>4.35</b>	<b>2</b>	<b>3.5</b>
<b>1. Goals and planning</b>	<b>4</b>	<b>4.27</b>	<b>1</b>	<b>3.6</b>
3. Social support	5	4.09	5	3.3
4. Shaping knowledge	6	4.03	3	3.3
11. Regulation	7	3.94	11	3.0
13. Identity	8	3.40	14	2.8
7. Associations	9	3.35	4	3.3
9. Comparison of outcomes	10	3.32	10	3.1
6. Comparison of behaviour	11	3.26	13	2.9
10. Reward and threat	12	3.21	7	3.2
14. Scheduled consequences	13	3.11	15	2.8
12. Antecedents	14	3.05	8	3.2
5. Natural consequences	15	2.94	16	2.7
16. Covert learning	16	2.33	9	3.1

Most important are denoted by bold text and most susceptible are denoted by italic text.

in more specific long-term relevant contexts shows that the importance ranking in Table 6 broadly aligns with prior works. For example, for physical activity and healthy eating Samdal et al. (2017) highlight feedback and monitoring, goals and planning, and social support (respectively ranked 3, 4, and 5 in Table 6). This is mirrored for smoking cessation by O’Neill et al. (2018) – highlighting feedback and monitoring – and for lifestyle changes by Willems et al. (2019) – highlighting goals and planning and repetition and substitution (ranked 2 in Table 6). Vestjens et al. (2015) also highlight repetition and substitution as generally important for complex interventions. Further, based on a review of electronic lifestyle activity

monitors (comparable to a sub-set of our real-world cases (Table 3)), Lyons et al. (2014) also identified goals and planning and feedback and monitoring as key BCTT clusters. This also aligns with good agreement between our expert sample and real-world cases (Table 3). In addition, Samdal et al. (2017) highlight the importance of person centredness and support for autonomy when maintaining behaviour over time, aligning with our identification of self-belief as the most important BCTT for long-term behavioural design. Overall, this suggests that while our results are both abstract and framed with respect to the long-term in general the distilled ranking is credible and comparable to other research on BCTT prioritisation in this area.

In terms of design, our results suggest that despite significant differences across behaviour change contexts there are some generally relevant clusters, which can help guide behavioural designers when identifying directions for exploration and development of potential interventions. Here, the abstract and generic nature of the BCTT clusters means that designers have significant scope for abduction and reframing in search of the most contextually relevant and valuable intervention (Cash, Valles Gamundi, et al. 2022b), whilst retaining the general explanative power offered by the BCTT. This also serves to elaborate the more general recognition that long-term behavioural design demands specific priorities as discussed by, for example, Fogg (2009b) or Tromp and Hekkert (2018). This illustrates how Table 6 can help designers to make sense of and explain their work in the face of the complex realities of long-term behavioural design and the associated challenges linked to designing interventions in this context (Schmidt and Stenger 2021). Thus, while such rankings cannot and should not straitjacket design exploration, they can serve an important role in guiding focus and identifying differences in short- and long-term behavioural design.

The second main contribution is based on taking a first step towards answering RQ2 (how susceptible are these BCTT clusters to change over time and in different contexts?) by highlighting those BCTT clusters most and least susceptible to change over time (based on averaged susceptibility results for each of the four factors; see Table 6) and qualitatively examining considerations relevant to long-term behavioural design (see Figure 2 and Table 5). Here, while there was general agreement on the need to consider change in context when applying BCTT clusters, there was less agreement on specific susceptibility and a high degree of variation in contextual embodiment. This is again aligned with expectations from prior research in design, which have highlighted how embodiment can vary significantly with respect to common underlying behaviour change techniques (Cash et al. 2020; Kelders et al. 2012). Overall, this suggests that despite diversity in responses our results are credible in comparison to prior research in this area (Bhamra, Lilley, and Tang 2011; Lockton, Harrison, and Stanton 2010; Tromp, Hekkert, and Verbeek 2011).

In terms of design, our results highlight how much of the variation in BCTT cluster prioritisation can be connected to the critical and substantial translation between abstract clusters and embodied interventions (Bhamra, Lilley, and Tang 2011; Tromp and Hekkert 2014). Notably, we describe how contextual embodied interventions vary substantially in form, often combine BCTT clusters, and even express the same cluster via multiple complementary aspects of the embodiment. This provides further evidence for the need to better understand how design and behavioural science guidance can be brought together to deliver impactful interventions that are also robust and explicable (Cash et al. 2022b; Mejía 2021; Schmidt 2020). Further, it highlights a major deficit in many of the behavioural-science led models of behavioural design where issues of contextual embodiment and

wider integration of design practices and insights are often minimised (Reid and Schmidt 2018; Schmidt and Stenger 2021). This is, for example, notable in the main process with which the BCTT clusters are associated – the behaviour change wheel – where embodiment issues are comparatively neglected (Michie, Atkins, and West 2015). More generally, while there are numerous social and behavioural theories dealing with aspects of behaviour in the long-term (such as self-determination) (Samdal et al. 2017), these are little connected to the design frameworks used to understand contextual, temporally dynamic, and embodied interventions (Pschetz and Bastian 2018). As such, our findings highlight that while important information is held in behavioural-science guidance – and this cannot be neglected – there is also a critical need to integrate design approaches, especially in the translation between abstract BCTT clusters and embodied interventions. Thus, while this work provides important insights linking temporal and context change to the robustness of behavioural design interventions – extending current work in design (Bay Brix Nielsen, Daalhuizen, and Cash 2021; Pschetz and Bastian 2018) – it also highlights the need for further research and theory development in this area.

### **Limitations and further work**

Before discussing implications, it is important to consider the main limitations of this work. First, while 12 experts is an acceptable sample for the Delphi method (and comparable to prior studies in both the BCTT and design contexts (McMahon and Bhamra 2015; Vestjens et al. 2015)), it places a focus on reaching consensus between diverse experts (Clayton 1997; Garvin and Simon 2017; Humphrey-Murto et al. 2020). Thus, we both recruited experts from across the spectrum of behavioural design and drew on real-world cases to support our conclusions. As such, our results provide a reasonable basis for prioritising the importance of BCTT clusters for long-term behavioural design but also highlight the need to better understand their contextualised embodiment in real-world interventions and the interaction between design and behavioural science guidance.

Second, while there is debate as to the required level of consensus for Delphi results, we followed prior guidance in considering BCT clusters to have reached consensus when percentage agreement was  $\geq 50\%$  (Gracht 2012). This highlighted a lack of consensus with respect to the susceptibility of BCTTs, despite the overall importance of this aspect being highlighted by all experts. Thus, while we provide several design relevant qualitative insights that help to elaborate this issue there is a critical need for further theory and empirical study linking behavioural science explanations of susceptibility to relevant design frameworks.

### **Implications**

Our findings hold two major implications for long-term behavioural design. First, designers in this area should be aware of the general differences evident when dealing with long- versus short-term behaviour change. For example, we highlight the prioritisation of self-belief, repetition and substitution, feedback and monitoring, and goals and planning, and deprioritise scheduled consequences, antecedents, natural consequences, and covert learning (see the importance ranking in Table 6). This helps designers in navigating the complexity of the BCTT clusters to identify potentially fruitful directions for exploration

but, critically, should not be considered a prescriptive straitjacket due to our second implication.

Second, designers should pay particular attention to the translation from abstract guidance to contextual, embodied intervention and be aware that this has the potential to significantly impact outcomes. Here, we highlight how numerous BCTT clusters are generically vulnerable to changing contextual factors (e.g. feedback and monitoring are considered particularly vulnerable to time and technological environment) as well as how successful interventions display a high degree of variety in BCTT embodiment, often combine BCTT clusters, or leveraging multiple embodiments of a single cluster in order to reinforce effects in context. In particular, the results in Table 5 provide an indicator of the design scope when developing interventions. However, care should be taken to ensure that the link between abstract explanation and embodied design are not lost. Our results empower designers in making the most of design guidance when developing interventions and also help in connecting these back to abstract techniques in order to balance the dual demands of efficacy and effectiveness characteristic of behavioural design.

## Conclusions

Long-term behaviour change is essential to addressing many societal and personal challenges, yet there is a critical gap in guidance for designers who must deliver long-term behavioural design. Hence, in this work we have sought to answer both: what are the most important Behaviour Change Technique Taxonomy (BCTT) clusters for long-term behavioural design (RQ1) and how susceptible are these BCTT clusters to change over time and in different contexts (RQ2)? We addressed these questions using a Delphi survey method with 12 international experts on behavioural intervention complemented by a reanalysis of over 100 real-world cases.

Based on our results, we provide a first general ranking of behaviour change techniques for long-term behavioural design and highlight how these are susceptible to change over time, as well as changes in the physical, social, and technological environment. Further, we elaborate the considerations associated with the translation from abstract BCTT clusters to embodied interventions and explore the scope of this design work. This provides essential and novel guidance for behavioural designers working in domains such as health, wellbeing, and sustainability, as well as contributing to wider research on how to deal with the demands of long-term behaviour change.

## Acknowledgements

The authors are grateful to the designers behind the real-world cases and the experts on the multidisciplinary Delphi panel who generously gave their time to participate in the studies.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

The work was co-funded by the Horizon 2020 ReHyb project under Grant agreement ID: 871767 and by the Horizon 2020 EURITO project under Grant agreement ID: 770420.

## ORCID

Agata Wrobel  <http://orcid.org/0000-0003-4441-0872>  
 Anja Maier  <http://orcid.org/0000-0002-3890-6452>  
 John Paulin Hansen  <http://orcid.org/0000-0001-5594-3645>  
 Philip Cash  <http://orcid.org/0000-0001-6498-0237>

## References

- Adams, M. K., and S. S. Meyer. 2009. "Use of a Modified Delphi Technique to Define Creative Scholarship and Identify Criteria to Evaluate Creative Scholarship in Textiles and Apparel Design in the United States." *The Design Journal* 12 (2): 131–52. <https://doi.org/10.2752/175630609X433102>.
- Bay Brix Nielsen, C. K. E., J. Daalhuizen, and P. Cash. 2021. "Defining the Behavioural Design Space." *International Journal of Design* 15 (1): 1–16. <http://www.ijdesign.org/index.php/IJDesign/article/view/3922>.
- Bhamra, T., D. Lilley, and T. Tang. 2011. "Design for Sustainable Behaviour: Using Products to Change Consumer Behaviour." *The Design Journal* 14 (4): 427–45. <https://doi.org/10.2752/175630611X13091688930453>. <http://www.ingentaconnect.com/content/bloomsbury/dsgj/2011/00000014/0000004/art00004>.
- Bohlen, L. C., S. Michie, M. de Bruin, A. J. Rothman, M. P. Kelly, H. N. K. Groarke, R. N. Carey, J. Hale, and M. Johnston. 2020. "Do Combinations of Behavior Change Techniques That Occur Frequently in Interventions Reflect Underlying Theory?" *Annals of Behavioral Medicine* 54 (11): 827–42. <https://doi.org/10.1093/abm/kaa078>.
- Bucher, A. 2020. *Engaged: Designing for Behavior Change*. New York, NY: Rosenfeld Media.
- Cash, P., C. Gram Hartlev, and C. Durazo. 2017. "Behavioural Design: A Process for Integrating Behaviour Change and Design." *Design Studies* 48 (January): 96–128. <https://doi.org/10.1016/j.destud.2016.10.001>.
- Cash, P., O. Isaksson, A. Maier, and J. D. Summers. 2022a. "Sampling in Design Research: Eight Key Considerations." *Design Studies* 78: 101077. <https://doi.org/10.1016/j.destud.2021.101077>.
- Cash, P., P. Khadilkar, J. Jensen, C. Dusterdich, and R. Mugge. 2020. "Designing Behaviour Change: A Behavioural Problem/Solution (BPS) Matrix." *International Journal of Design* 14 (2): 65–83. <http://www.ijdesign.org/index.php/IJDesign/article/view/3952>.
- Cash, P., X. Valles Gamundi, I. Echstrom, and J. Daalhuizen. 2022b. "Method Use in Behavioural Design: What, How, and Why?" *International Journal of Design* 16 (1): 1–21. <http://www.ijdesign.org/index.php/IJDesign/article/view/4352/973>.
- Clayton, M. J. 1997. "Delphi: A Technique to Harness Expert Opinion for Critical Decision-Making Tasks in Education." *Educational Psychology* 17 (4): 373–86. <https://doi.org/10.1080/0144341970170401>.
- Davies, E., J. Martin, and D. Foxcroft. 2016. "Development of an Adolescent Alcohol Misuse Intervention Based on the Prototype Willingness Model: A Delphi Study." *Health Education* 116 (3): 275–91. <https://doi.org/10.1108/HE-01-2015-0006>.
- Dolan, P., M. Hallsworth, D. Halpern, D. King, and I. Vlaev. 2014. "MindSpace: Influencing Behaviour through Public Policy." <http://elibrary.kiu.ac.ug:8080/jspui/handle/1/1520>.
- Dorst, K. 2011. "The Core of 'Design Thinking' and Its Application." *Design Studies* 32 (6): 521–32. <https://doi.org/10.1016/j.destud.2011.07.006>.
- Fogg, B. J. 2009a. "Creating Persuasive Technologies: An Eight-Step Design Process." In *Persuasive Technology. Fourth International Conference, PERSUASIVE 2009*, 1–6. Claremont, CA, USA. <http://itp.nyu.edu/sd107/diyhealth2012/wp-content/uploads/2012/09/Fogg-2009-Creating-persuasive-technologies-an-eight-step-design-process.pdf>.
- Fogg, B. J. 2009b. "The Behavior Grid: 35 Ways Behavior Can Change." In *Proceedings of the 4th International Conference on Persuasive Technology*, 1–5. <http://dl.acm.org/citation.cfm?id=1542001>.
- Garvin, L. A., and S. R. Simon. 2017. "Prioritizing Measures of Digital Patient Engagement: A Delphi Expert Panel Study." *Journal of Medical Internet Research* 19 (5): 1–6. <https://doi.org/10.2196/JMIR.4778>.

- Gourlan, M., P. Bernard, C. Bortolon, A. J. Romain, O. Lareyre, M. Carayol, G. Ninot, and J. Boiché. 2016. "Efficacy of Theory-Based Interventions to Promote Physical Activity. A Meta-Analysis of Randomised Controlled Trials." *Health Psychology Review* 10 (1): 50–66. <https://doi.org/10.1080/17437199.2014.981777>.
- Gracht, H. A. 2012. "Consensus Measurement in Delphi Studies. Review and Implications for Future Quality Assurance." *Technological Forecasting and Social Change* 79 (8): 1525–36. <https://doi.org/10.1016/j.techfore.2012.04.013>.
- Green, K., and P. Vergragt. 2002. "Towards Sustainable Households: A Methodology for Developing Sustainable Technological and Social Innovations." *Futures-the Journal of Forecasting Planning and Policy* 34 (5): 381–400. [https://doi.org/10.1016/S0016-3287\(01\)00066-0](https://doi.org/10.1016/S0016-3287(01)00066-0).
- Hagger, M. S., and M. Weed. 2019. "DEBATE: Do Interventions Based on Behavioral Theory Work in the Real World?" *International Journal of Behavioral Nutrition and Physical Activity* 16 (36): 1–10. <https://doi.org/10.1186/s12966-019-0795-4>.
- Hedin, B., C. Katzeff, E. Eriksson, and D. Pargman. 2019. "A Systematic Review of Digital Behaviour Change Interventions for More Sustainable Food Consumption." *Sustainability*, <https://doi.org/10.3390/su11092638>.
- Humphrey-Murto, S., T. J. Wood, C. Gonsalves, K. Mascioli, and L. Varpio. 2020. "The Delphi Method." *Academic Medicine: Journal of the Association of American Medical Colleges* 95 (1): 168. <https://doi.org/10.1097/ACM.0000000000002887>.
- Jack, K., S. M. McLean, J. K. Moffett, and E. Gardiner. 2010. "Barriers to Treatment Adherence in Physiotherapy Outpatient Clinics: A Systematic Review." *Manual Therapy*, <https://doi.org/10.1016/j.math.2009.12.004>.
- Kelders, S. M., R. N. Kok, H. C. Ossebaard, and J. van Gemert-Pijnen. 2012. "Persuasive System Design Does Matter: A Systematic Review of Adherence to Web-Based Interventions." *Journal of Medical Internet Research* 14 (6): e152. <https://doi.org/10.2196/jmir.2104>.
- Kelly, M. P., and M. Barker. 2016. "Why is Changing Health-Related Behaviour So Difficult?" *Public Health* 136: 109–16. <https://doi.org/10.1016/j.puhe.2016.03.030>.
- Khadilkar, P., and P. Cash. 2020. "Understanding Behavioural Design: Barriers and Enablers." *Journal of Engineering Design* 31 (10): 508–29. <https://doi.org/10.1080/09544828.2020.1836611>.
- Kouprie, M., and F. S. Visser. 2009. "A Framework for Empathy in Design: Stepping Into and Out of the User's Life." *Journal of Engineering Design* 20 (5): 437–48. <https://doi.org/10.1080/09544820902875033>.
- Kwasnicka, D., S. U. Dombrowski, M. White, and F. Sniehotta. 2016. "Theoretical Explanations for Maintenance of Behaviour Change: A Systematic Review of Behaviour Theories." *Health Psychology Review*, <https://doi.org/10.1080/17437199.2016.1151372>.
- Lally, P., C. H. M. van Jaarsveld, H. W. W. Potts, and J. Wardle. 2010. "How Are Habits Formed: Modelling Habit Formation in the Real World." *European Journal of Social Psychology* 40 (6): 998–1009. <https://doi.org/10.1002/ejsp.674>.
- Lilley, D., and G. T. Wilson. 2013. "Integrating Ethics into Design for Sustainable Behaviour." *Journal of Design Research* 11 (3): 278–99. <https://doi.org/10.1504/JDR.2013.056593>.
- Lockton, D., D. Harrison, and N. Stanton. 2008. "Making the User More Efficient: Design for Sustainable Behaviour." *International Journal of Sustainable Engineering* 1 (1): 3–8. <https://doi.org/10.1080/19397030802131068>.
- Lockton, D., D. Harrison, and N. A. Stanton. 2010. "The Design with Intent Method: A Design Tool for Influencing User Behaviour." *Applied Ergonomics* 41 (3): 382–92. <https://doi.org/10.1016/j.apergo.2009.09.001>.
- Lyons, E. J., Z. H. Lewis, B. G. Mayrsohn, and J. L. Rowland. 2014. "Behavior Change Techniques Implemented in Electronic Lifestyle Activity Monitors: A Systematic Content Analysis." *Journal of Medical Internet Research* 16 (8): e192. <https://doi.org/10.2196/jmir.3469>.
- MacPherson, M. M., K. D. Cranston, S. R. Locke, J. E. Bourne, and M. E. Jung. 2021. "Using the Behavior Change Wheel to Develop Text Messages to Promote Diet and Physical Activity Adherence Following a Diabetes Prevention Program." *Translational Behavioral Medicine* 11 (8): 1585–95. <https://doi.org/10.1093/tbm/ibab058>.

- Maier, A., and P. Cash. 2022. "Designing for Human Behaviour in a Systemic World." In *Handbook of Engineering Systems Design*, 1–34. Springer International Publishing. [https://doi.org/10.1007/978-3-030-46054-9\\_16-1](https://doi.org/10.1007/978-3-030-46054-9_16-1)
- McCarthy, M. J., Y. E. Garcia, D. J. Dunn, K. S. Lyons, and T. Bakas. 2020. "Topics in Stroke Rehabilitation Development and Validation of a Quality of Relationship Intervention for Stroke Survivor-Family Caregiver Dyads." *Topics in Stroke Rehabilitation* 27 (4): 305–15. <https://doi.org/10.1080/10749357.2019.1690823>.
- McMahon, M., and T. Bhamra. 2015. "Social Sustainability in Design: Moving the Discussions Forward." *The Design Journal* 18 (3): 367–91. <https://doi.org/10.1080/14606925.2015.1059604>.
- Mejía, G. M. 2021. "Theory-Driven or Theory-Informed? A Review of Behavioural Economics in Design." *The Design Journal* 24 (4): 567–87. <https://doi.org/10.1080/14606925.2021.1935089>.
- Meng, K., B. Seekatz, G. Haug, G. Mosler, B. Schwaab, U. Worringer, and H. Faller. 2014. "Evaluation of a Standardized Patient Education Program for Inpatient Cardiac Rehabilitation: Impact on Illness Knowledge and Self-Management Behaviors up to 1 Year." *Health Education Research* 29 (2): 235–46. <https://doi.org/10.1093/her/cyt107>.
- Michie, S., L. Atkins, and R. West. 2015. *The Behaviour Change Wheel: A Guide to Designing Interventions*. London: Silverback Publishing.
- Michie, S., M. Richardson, M. Johnston, C. Abraham, J. Francis, W. Hardeman, M. P. Eccles, J. Cane, and C. E. Wood. 2013. "The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions." *Annals of Behavioral Medicine* 46 (1): 81–95. <https://doi.org/10.1007/s12160-013-9486-6>.
- Niedderer, K., S. Clune, and G. Ludden. 2017. *Design for Behaviour Change: Theories and Practices of Designing for Change*. Design for Behaviour Change: Theories and Practices of Designing for Change. Taylor and Francis. <https://doi.org/10.4324/9781315576602>.
- Niedderer, K., G. Ludden, S. Clune, D. Lockton, J. Mackrill, A. Morris, R. Cain, et al. 2016. "Design for Behaviour Change as a Driver for Sustainable Innovation: Challenges and Opportunities for Implementation in the Private and Public Sectors." *International Journal of Design* 10 (2): 67–85. <http://www.ijdesign.org/index.php/IJDesign/article/view/2260>.
- Nieuwenhuijsen, E. R., E. Zemper, K. R. Miner, and M. Epstein. 2006. "Health Behavior Change Models and Theories: Contributions to Rehabilitation." *Disability and Rehabilitation* 28 (5): 245–56. <https://doi.org/10.1080/09638280500197743>.
- Okoli, C., and S. D. Pawlowski. 2004. "The Delphi Method as a Research Tool: An Example, Design Considerations and Applications." *Information & Management* 42 (1): 15–29. <https://doi.org/10.1016/j.im.2003.11.002>.
- O'Neill, N., O. Dogar, M. Jawad, I. Kellar, M. Kanaan, and K. Siddiqi. 2018. "Which Behavior Change Techniques May Help Waterpipe Smokers to Quit? An Expert Consensus Using a Modified Delphi Technique." *Nicotine & Tobacco Research* 20 (2): 154–60. <https://doi.org/10.1093/ntr/ntw297>.
- Onwuegbuzie, A. J., and N. L. Leech. 2007. "Sampling Designs in Qualitative Research: Making the Sampling Process More Public." *The Qualitative Report* 12 (2): 19–20. <https://doi.org/10.1007/s11135-006-9000-3>.
- Pschetz, L., and M. Bastian. 2018. "Temporal Design: Rethinking Time in Design." *Design Studies* 56: 169–84. <https://doi.org/10.1016/j.destud.2017.10.007>.
- Reid, S., and R. Schmidt. 2018. "A New Model for Integrating Behavioral Science and Design." *Behavioral Scientist*. <http://behavioralscientist.org/a-new-model-for-integrating-behavioral-science-and-design/>.
- Rowe, G., G. Wright, and F. Bolger. 1991. "Delphi: A Reevaluation of Research and Theory." *Technological Forecasting and Social Change* 39 (3): 235–51. [https://doi.org/10.1016/0040-1625\(91\)90039-l](https://doi.org/10.1016/0040-1625(91)90039-l).
- Samdal, G. B., G. E. Eide, T. Barth, G. Williams, and E. Meland. 2017. "Effective Behaviour Change Techniques for Physical Activity and Healthy Eating in Overweight and Obese Adults; Systematic Review and Meta-Regression Analyses." *International Journal of Behavioral Nutrition and Physical Activity*, <https://doi.org/10.1186/s12966-017-0494-y>.
- Schmidt, R. 2020. "Strange Bedfellows: Design Research and Behavioral Design." *DRS2020: Synergy* 1: 11–14. <https://doi.org/10.21606/drs.2020.252>.

- Schmidt, R. 2022. "A Model for Choice Infrastructure: Looking Beyond Choice Architecture in Behavioral Public Policy." *Behavioural Public Policy*, 1–26. <https://doi.org/10.1017/bpp.2021.44>.
- Schmidt, R., and K. Stenger. 2021. "Behavioral Brittleness: The Case for Strategic Behavioral Public Policy." *Behavioural Public Policy*, 1–26. <https://doi.org/10.1017/bpp.2021.16>.
- Sumsion, T. 1998. "The Delphi Technique: An Adaptive Research Tool." *British Journal of Occupational Therapy* 61 (4): 153–56. <https://doi.org/10.1177/030802269806100403>.
- Teasell, R., S. Mehta, S. Pereira, A. McIntyre, S. Janzen, L. Allen, L. Lobo, and R. Viana. 2012. "Time to Rethink Long-Term Rehabilitation Management of Stroke Patients." *Topics in Stroke Rehabilitation* 19 (6): 457–62. <https://doi.org/10.1310/tsr1906-457>.
- Tromp, N., and P. Hekkert. 2014. "Social Implication Design (SID) – A Design Method to Exploit the Unique Value of the Artefact to Counteract Social Problems." In *Proceedings of the DRS 2014*, 1–16. Umea: Umea Institute of Design.
- Tromp, N., and P. Hekkert. 2018. *Designing for Society: Products and Services for a Better World*. Bloomsbury Visual Arts. [https://books.google.dk/books?hl=en&lr=&id=o6dvDwAAQBAJ&oi=fnd&pg=PP1&dq=info:Il-dVRZvHtwJ:scholar.google.com&ots=iDTJH\\_Yt6l&sig=yDEndTuH6BGGtboijqsy1yHdkBM&redir\\_esc=y#v=onepage&q&f=false](https://books.google.dk/books?hl=en&lr=&id=o6dvDwAAQBAJ&oi=fnd&pg=PP1&dq=info:Il-dVRZvHtwJ:scholar.google.com&ots=iDTJH_Yt6l&sig=yDEndTuH6BGGtboijqsy1yHdkBM&redir_esc=y#v=onepage&q&f=false).
- Tromp, N., P. Hekkert, and P.-P. P. Verbeek. 2011. "Design for Socially Responsible Behavior: A Classification of Influence Based on Intended User Experience." *Design Issues* 27 (3): 3–19. [https://doi.org/10.1162/DESI\\_a\\_00087](https://doi.org/10.1162/DESI_a_00087).
- Ubhi, H. K., S. Michie, D. Kotz, O. C. P. van Schayck, A. Selladurai, and R. West. 2016. "Characterising Smoking Cessation Smartphone Applications in Terms of Behaviour Change Techniques, Engagement and Ease-of-Use Features." *Translational Behavioral Medicine* 6 (3): 410–17. <https://doi.org/10.1007/s13142-015-0352-x>.
- Vestjens, L., G. I. J. M. Kempen, R. Crutzen, G. Kok, and G. A. R. Zijlstra. 2015. "Promising Behavior Change Techniques in a Multicomponent Intervention to Reduce Concerns About Falls in Old Age: A Delphi Study." *Health Education Research* 30 (2): 309–22. <https://doi.org/10.1093/her/cyv003>.
- Walsh, K. A., S. Timmons, S. Byrne, J. Browne, and J. Mc Sharry. 2021. "Identifying Behavior Change Techniques for Inclusion in a Complex Intervention Targeting Antipsychotic Prescribing to Nursing Home Residents with Dementia." *Translational Behavioral Medicine* 11 (2): 470–83. <https://doi.org/10.1093/tbm/ibaa053>.
- Watson, J., O. Cumming, A. MacDougall, A. Czerniewska, and R. Dreibelbis. 2021. "Effectiveness of Behaviour Change Techniques Used in Hand Hygiene Interventions Targeting Older Children – A Systematic Review." *Social Science & Medicine*, <https://doi.org/10.1016/j.socscimed.2021.114090>.
- Wendel, S. 2013. *Designing for Behavior Change: Applying Psychology and Behavioral Economics*. O'Reilly. <https://books.google.com/books?hl=en&lr=&id= SXH0AQAAQBAJ&oi=fnd&pg=PR3&dq=%22designing+for+behavior+change%22&ots=Lqk4yX8Rcu&sig=ks2XXOyulVjpT-zT9wNBe6581gc>.
- West, R., and S. Michie. 2016. *A Guide to Development and Evaluation of Digital Interventions in Healthcare*. London: Silverback Publishing.
- Willems, M., A. Waninge, J. de Jong, T. I. M. Hilgenkamp, and C. P. van der Schans. 2019. "Exploration of Suitable Behaviour Change Techniques for Lifestyle Change in Individuals with Mild Intellectual Disabilities: A Delphi Study." *Journal of Applied Research in Intellectual Disabilities* 32 (3): 543–57. <https://doi.org/10.1111/jar.12548>.
- Williams, S., and D. French. 2011. "What Are the Most Effective Intervention Techniques for Changing Physical Activity Self-Efficacy and Physical Activity Behaviour – And Are They the Same?" *Health Education Research* 26 (2): 308–22. <https://doi.org/10.1093/her/cyr005>.



## Appendix

**Table 7.** Detail of consensus results for time (Question 2).

BCT	Sample			
	Design		Behavioural Science	
	Consensus	Median	Consensus	Median
1. Goals and planning	67	3.0	50	3.5
2. Feedback and monitoring	33	3.5	33	3.0
3. Social support	33	3.0	50	3.0
4. Shaping knowledge	50	3.0	50	3.5
5. Natural consequences	33	3.5	50	3.0
6. Comparison of behaviour	50	2.0	50	3.0
7. Associations	50	2.5	50	3.0
8. Repetition and substitution	67	2.0	50	4.0
9. Comparison of outcomes	67	2.0	50	2.5
10. Reward and threat	50	4.5	67	4.0
11. Regulation	33	3.0	33	3.0
12. Antecedents	50	2.0	50	3.0
13. Identity	67	2.0	50	3.5
14. Scheduled consequences	83	2.0	33	2.5
15. Self-belief	50	2.0	50	3.5
16. Covert learning	50	2.5	67	3.0

**Table 8.** Detail of design consensus results for Question 3.

BCT	Environment					
	3a physical		3b social		3c technological	
	Consensus	Median	Consensus	Median	Consensus	Median
1. Goals and planning	33	4.0	67	4.0	83	4.0
2. Feedback and monitoring	50	4.0	33	3.5	50	3.5
3. Social support	50	4.0	67	4.0	67	2.0
4. Shaping knowledge	67	4.0	50	3.5	50	3.5
5. Natural consequences	33	3.0	33	3.5	50	2.0
6. Comparison of behaviour	50	2.5	33	3.5	50	2.5
7. Associations	50	4.0	50	3.5	50	3.5
8. Repetition and substitution	50	3.0	67	4.0	50	3.0
9. Comparison of outcomes	67	4.0	67	3.0	50	4.0
10. Reward and threat	50	3.0	50	3.5	50	3.0
11. Regulation	50	4.0	50	2.5	67	2.0
12. Antecedents	33	3.0	67	3.0	50	2.5
13. Identity	33	2.5	50	3.5	50	2.5
14. Scheduled consequences	67	3.0	50	2.5	50	3.0
15. Self-belief	50	4.0	67	4.0	50	2.5
16. Covert learning	67	3.0	50	3.0	50	3.0

**Table 9.** Detail of behavioural science consensus results for Question 3.

BCT	Environment					
	3a physical		3b social		3c technological	
	Consensus	Median	Consensus	Median	Consensus	Median
1. Goals and planning	50	3.0	50	4.0	50	3.0
2. Feedback and monitoring	67	3.0	67	4.0	50	4.0
3. Social support	50	2.5	67	5.0	67	2.0
4. Shaping knowledge	50	2.5	83	4.0	50	3.5
5. Natural consequences	67	3.0	50	3.0	50	1.5
6. Comparison of behaviour	50	2.5	67	4.0	50	2.0
7. Associations	33	4.0	50	2.5	33	3.0
8. Repetition and substitution	50	4.0	50	3.5	50	3.5
9. Comparison of outcomes	50	3.5	67	4.0	67	2.0
10. Reward and threat	50	4.0	50	3.5	33	2.5
11. Regulation	50	3.5	67	3.0	50	2.5
12. Antecedents	67	4.0	50	3.5	50	3.5
13. Identity	50	2.0	67	4.0	50	3.0
14. Scheduled consequences	50	3.0	50	3.5	50	2.0
15. Self-belief	50	3.0	67	4.0	67	2.0
16. Covert learning	50	3.5	50	4.0	83	3.0

**Table 10.** Detail of ranking of importance for long-term behaviour change.

BCT	Sub-sample					
	Design		Behavioural Science		Cases	
	Rank	Mean	Rank	Mean	Rank	Mean
1. Goals and planning	5	3.8	1	4.5	5	4.5
2. Feedback and monitoring	2	4.3	4	4.2	4	4.5
3. Social support	4	4.0	4	4.2	7	4.1
4. Shaping knowledge	7	3.7	3	4.3	8	4.1
5. Natural consequences	11	3.0	14	3.0	13	2.8
6. Comparison of behaviour	16	2.5	9	3.2	6	4.1
7. Associations	5	3.8	14	3.0	11	3.2
8. Repetition and substitution	1	4.5	7	4.0	1	5.0
9. Comparison of outcomes	11	3.0	9	3.2	9	3.8
10. Reward and threat	11	3.0	9	3.2	10	3.5
11. Regulation	11	3.0	8	3.8	1	5.0
12. Antecedents	7	3.7	9	3.2	15	2.3
13. Identity	10	3.3	4	4.2	14	2.7
14. Scheduled consequences	9	3.5	16	2.8	12	3.0
15. Self-belief	2	4.3	1	4.5	1	5.0
16. Covert learning	15	2.8	9	3.2	16	1.0

**Table 11.** Detail of cases mean calculation.

BCT	Interventions		Normalised	
	Total	Long-term	Score	1–5 Mean
1. Goals and planning	23	20	0.9	4.5
2. Feedback and monitoring	44	39	0.9	4.5
3. Social support	9	7	0.8	4.1
4. Shaping knowledge	22	17	0.8	4.1
5. Natural consequences	11	5	0.5	2.8
6. Comparison of behaviour	32	25	0.8	4.1
7. Associations	56	31	0.6	3.2
8. Repetition and substitution	5	5	1.0	5.0
9. Comparison of outcomes	10	7	0.7	3.8
10. Reward and threat	13	8	0.6	3.5
11. Regulation	1	1	1.0	5.0
12. Antecedents	82	27	0.3	2.3
13. Identity	7	3	0.4	2.7
14. Scheduled consequences	2	1	0.5	3.0
15. Self-belief	2	2	1.0	5.0
16. Covert learning	2	0	0.0	1.0

**Table 12.** Means for susceptibility to other factors (Questions 2–3). Sample denoted by Design (D) and Behavioural Science (B).

BCT	Time		BCT	Physical environment	
	D Mean	B Mean		D Mean	B Mean
1. Goals and planning	3.5	3.2	1. Goals and planning	4.0	3.3
2. Feedback and monitoring	3.3	3.0	2. Feedback and monitoring	3.7	3.0
3. Social support	2.8	3.2	3. Social support	3.7	2.8
4. Shaping knowledge	2.8	3.2	4. Shaping knowledge	3.5	2.8
5. Natural consequences	3.0	2.5	5. Natural consequences	3.0	2.5
6. Comparison of behaviour	2.3	3.0	6. Comparison of behaviour	2.5	2.7
7. Associations	2.7	2.5	7. Associations	3.8	4.0
8. Repetition and substitution	2.7	3.8	8. Repetition and substitution	3.2	3.8
9. Comparison of outcomes	2.2	2.7	9. Comparison of outcomes	4.0	3.3
10. Reward and threat	4.0	3.5	10. Reward and threat	2.8	3.7
11. Regulation	2.8	3.2	11. Regulation	3.7	3.7
12. Antecedents	2.2	3.3	12. Antecedents	3.2	4.3
13. Identity	2.5	3.0	13. Identity	2.5	2.3
14. Scheduled consequences	2.2	2.5	14. Scheduled consequences	3.0	2.8
15. Self-belief	2.2	3.0	15. Self-belief	3.7	2.8
16. Covert learning	2.8	2.8	16. Covert learning	3.0	3.2

**Table 13.** Means for susceptibility to other factors (Questions 2–3).

BCT	Social environment		BCT	Technical environment	
	D Mean	B Mean		D Mean	B Mean
1. Goals and planning	3.5	4.2	1. Goals and planning	3.8	3.2
2. Feedback and monitoring	3.5	3.5	2. Feedback and monitoring	3.8	4.2
3. Social support	4.3	4.5	3. Social support	2.3	2.3
4. Shaping knowledge	3.7	3.8	4. Shaping knowledge	3.3	3.2
5. Natural consequences	3.3	2.8	5. Natural consequences	2.3	1.8
6. Comparison of behaviour	3.7	4.0	6. Comparison of behaviour	2.8	2.3
7. Associations	3.8	3.0	7. Associations	3.3	3.0
8. Repetition and substitution	3.3	3.3	8. Repetition and substitution	2.8	3.0
9. Comparison of outcomes	3.5	3.5	9. Comparison of outcomes	3.3	2.0
10. Reward and threat	3.5	3.3	10. Reward and threat	2.7	2.3
11. Regulation	2.8	3.0	11. Regulation	2.2	2.3
12. Antecedents	3.2	3.2	12. Antecedents	2.7	3.3
13. Identity	3.3	4.0	13. Identity	2.3	2.5
14. Scheduled consequences	3.0	3.2	14. Scheduled consequences	3.0	2.3
15. Self-belief	3.3	4.0	15. Self-belief	2.5	2.0
16. Covert learning	3.3	3.7	16. Covert learning	3.0	2.8