Development of a formal system for representing behaviour change theories

Robert West[1*]
Cristina A. Godinho[2,3]
Lauren Connell Bohlen[2,4]
Rachel N. Carey[2,5]
Janna Hastings[6]
Carmen E. Lefevre[2,7]
Susan Michie[2]

[1] Department of Behavioural Science and Health, University College London, WC1E 7HB, UK
[2] Centre for Behaviour Change, University College London, WC1E 7HB, UK
[3] Instituto Universitário de Lisboa (ISCTE-IUL), CIS-IUL, Lisboa, Portugal
[4] Department of Kinesiology, University of Rhode Island, Kingston, RI, USA, 02881
[5] Zinc, 23 Tileyard Studios, Tileyard Road, London N7 9AH.
[6] Department of Biological Sciences, University of Cambridge, CB2 1RX, UK
[7] Healthbridge Ltd, London, EC1V 1NY

[*]Corresponding author

Abstract

Use of natural language to represent behaviour change theories has resulted in lack of clarity and consistency hindering comparison, integration, development and use. This paper describes development of a formal system for representing behaviour change theories which aims to improve clarity and consistency. A given theory is represented in terms of 1) its component constructs (e.g., 'self-efficacy', 'perceived threat', 'subjective norm') which are labelled and defined, and 2) relationships between pairs of constructs, which may be causal, structural, or semantic. This formalism appears adequate to represent five commonly used theories (Health Belief Model, Information-Motivation-Behavioural Skill Model, Social Cognitive Theory, Theory of Planned Behaviour, and the Transtheoretical Model). Theory authors and experts judged that the system was able to capture the main propositions of the theories. Following this proof-of-concept, the next step is to assess how far the system can be applied to other theories of behaviour change.

Keywords: Behaviour change, theories, theoretical construct, definitions, theory representation, methodology.

Behaviour change lies at the heart of many challenges facing society, such as improving public health and environmental sustainability [1,2]. Over the last century, a wide range of theories have been developed seeking to explain the processes underlying behaviour change for use in behaviour change interventions [3]. There is a need for theories to be represented in a more consistent and less ambiguous way to allow better comparison, integration, development and use. This paper describes initial development and evaluation of a formal system designed to achieve this.

The term 'theory' has been defined in many ways [4]. One definition agreed by a multi-disciplinary panel of experts from psychology, sociology, anthropology and economics is 'a set of concepts and/or statements which specify how phenomena relate to each other, providing an organising description of a system that accounts for what is known, and explains and predicts phenomena' [3]. In this paper we use the term 'constructs' rather than 'concepts'. Constructs are representations of things that are believed to exist in the world, including objects, processes and their attributes.

Theories can benefit scientific investigation and its application in several ways. They can summarise the current state of knowledge, broaden understanding, and stimulate the generation of new knowledge. They can help to structure thinking and guide research, providing a framework that facilitates communication across research groups [5]. Importantly, theories can provide a basis for intervention development [4,6,7], and evaluations of theory-based interventions can provide empirical tests of theories [8].

Reviews of behaviour change theories have noted that they are often not clearly specified nor capable of generating clear and accurate predictions [9-13]. Moreover, the informal ways in which they tend to be presented limit direct comparison and integration [14,15]. Advancing the science of behaviour change, and the application of that science, requires us to be able to compare theories in terms of their content, scope and predictions as a basis for selecting, integrating and modifying them.

Possibly as a result of this, intervention development has been largely disconnected from explicit theories. A majority of behaviour change interventions reported in published evaluations either make no reference to theory or apply it partially and inconsistently [15-17]. Clearer specification of these theories should increase their usefulness in intervention development and evaluation.

A review of 83 behaviour change theories revealed several potential areas for improvement [3,14]. First, where theories overlap in scope it is important to be clearer about the precise differences between them and the reasons for this. Secondly, it is important for theories to be clear about why they only include a limited subset of the constructs that appear to be relevant. Thirdly, theories are expressed using natural language, sometimes supplemented by diagrams that are constructed ad-hoc without a clear indication of what all the components are intended to represent. Natural language is efficient and highly expressive but introduces ambiguity by virtue of uncertain meaning of terms and a heavy reliance on context to disambiguate them. Different labels are used for the same construct even within a given theory; different theories use different labels for the same construct or the same label for different constructs.

A number of frameworks and resources have been used to organise the large number of constructs in the behaviour change literature. One approach is exemplified by the US National Institute of Health's 'Grid Enabled Measures' (GEM) web-based database. GEM provides descriptions of constructs, as well as measures for their assessment (https://www.gem-beta.org/Public/Home.aspx). However, it does not seek to compare or integrate constructs. A second approach has been to synthesise constructs from different theories. The Science of

Behaviour Change (SOBC) project is building a repository of measures of 'mechanisms of action' (https:scienceofbehaviorchange.org/measures). The Theoretical Domains Framework (TDF) identified 128 constructs from 33 theories of behaviour and integrated these into a framework, consisting of 14 theoretical domains [12,18]. While valuable, the GEM repository, the SOBC project and the TDF have not attempted to improve or standardise how theories are represented.

We aimed to develop, and undertake initial evaluation of, a formal system for representing theories of behaviour change that would improve their clarity and facilitate their comparison, integration, development and use. We did this by taking five commonly used theories and attempting to represent them using such a formalism. We then sought to assess their faithfulness to the theories through discussion with theory authors and experts. It is important to note that the goal was to represent the theories as accurately as possible and not to evaluate them. The next step would be to establish whether this formal system could be extended to other behaviour change theories.

Results

Constructs and their definitions

Table 1 shows the labels and definitions of main constructs identified in the five theories.

A total of 85 constructs were identified but these were not necessarily distinct; for example the Information-Motivation-Behavioural Skills model uses constructs from the Theory of Planned Behaviour. Some constructs were collections of other constructs (e.g., 'demographic variables' is a construct that includes attributes such as 'age', 'sex', etc.). Some of the constructs are attributes of individuals that are potentially modifiable (e.g., 'perceived susceptibility'), while others are modifiable attributes of the environment (e.g., 'cues to action'), and others are unmodifiable or structural attributes (e.g., 'age').

Table 1: Construct labels and definitions for five theories of behaviour change

Theory[1]	**Construct Label**	**Construct Definition**
HBM	1. Demographic variables	Variables representing demographic attributes of people, such as age and gender
HBM	2. Socio-psychological variables	Variables representing social, emotional, cognitive and behavioural attributes of people
HBM	3. Structural variables	Variables that affect people's perceptions and thus indirectly influence health-related behaviour (e.g. knowledge about a disease, prior experience of a disease)
HBM	4. Perceived threat	Perception of the personal threat posed by a particular disease or health problem
HBM	5. Perceived susceptibility	Perception of the personal risk of contracting a particular disease or developing a particular health problem
HBM	6. Perceived severity	Perception of the seriousness of the disease and associated emotional arousal elicited by thoughts of the disease and the anticipated difficulties that could be caused by the disease
HBM	7. Perceived benefits	Beliefs about the relative effectiveness of specified options for reducing a health threat, distinct from objective facts
HBM	8. Perceived barriers	Beliefs about the negative attributes of health-protective action (e.g., inconvenience, expense, discomfort)
HBM	9. Self-efficacy	Beliefs about how far one is capable of taking a health-protective action

HBM	10. Cues to action	Events that may elicit a health-protective action, which can be either internal or external
HBM	11. Likelihood of a health-protective action	The probability of taking action designed to prevent or mitigate threat to health
IMB	1. Enactment of AIDS preventive behaviour	Specific AIDS risk-reducing acts
IMB	2. Complexity of AIDS preventive behaviour	An attribute of specific AIDS risk-reducing acts which means that it requires complex behavioural skills
IMB	3. Performance of behavioural Skills	Objective and perceived abilities or proficiencies acquired through training and practice; the mechanisms through which information and motivation generally exert an influence on behaviour
IMB	4. Information[2]	Knowledge about how HIV is transmitted and prevented
IMB	5. Motivation	The impetus that gives purpose or direction to initiate and maintain behaviours (e.g., preventative behaviours). Comprised of attitudes, social norms, and their interrelations consistent with their specification in the TPB and Theory of Reasoned Action.
IMB	6. Information about means of transmission	Specific knowledge about how HIV is transmitted
IMB	7. Information about means of prevention	Specific knowledge about how HIV is prevented
IMB	8. Decision rules about transmission and prevention	Heuristic strategies for problem solving/decision-making that provide an efficient means for making somewhat automatic decisions regarding the behaviour
IMB	9. Implicit theories about risk and prevention	More complex sets of beliefs that require cognitive effort to apply in safer sex decisions,
IMB	10. Attitudes[2]	Individuals' favourable to unfavourable evaluation of personally performing a specified HIV preventive behaviour within a given time frame and context. The basic psychological underpinnings of attitudes towards an HIV preventive act involve the sum of beliefs about the consequences of a behaviour multiplied by evaluations of the consequences of the behaviour (intended to be the same as in TPB).
IMB	11. Social norms	Individuals' perception what significant referent others wish the individual to do with respect to performing a specified HIV preventive behaviour within a given time frame and context (intended to be the same as in TPB)
IMB	12. Self-efficacy	Self-belief in one's ability to use one's skills effectively to carry out behaviour (intended to be the same as in HBM)
IMB	13. Ability to practise self and partner-reinforcement	The ability to reinforce oneself or one's sexual partner's safer sex practices
IMB	14. Group-specific skills	Prevention relevant behavioural skills that are pertinent only to certain groups of people (E.g. substance abuse status, gender, ethnicity)
IMB	15. Self-acceptance of sexuality	The ability to acknowledge that one is sexually active
IMB	16. Ability to acquire behaviourally relevant information	The ability to obtain accurate information about methods of preventing HIV that can be readily translated into preventive behaviour by an individual
IMB	17. Ability to negotiate AIDS prevention with partner	The ability to discuss and agree upon AIDS-preventative actions with a sexual partner, and to remove oneself from situations in which safer sex cannot be agreed upon
IMB	18. Ability to practise public prevention acts	The ability to carry out disease relevant protective behaviours that must be performed out in the open or not in

		completely private settings (e.g. purchasing condoms and undergoing HIV testing)
IMB	19. Consistent AIDS prevention ability	The ability to routinely engage in HIV protective behaviours over time and situations
IMB	20. Beliefs about outcomes	Beliefs about the consequences of a behaviour
IMB	21. Evaluations of outcomes	Evaluations of the consequences of a behaviour
IMB	22. Beliefs about support/opposition from significant others	Having important referent others who support or oppose one's practice of a particular behaviour
IMB	23. Motivation to comply with referent others' wishes	Willingness to act in accordance with referent others' opinions (intended to be the same as TPB)
SCT	1. Behaviour	Any action one takes in response to internal/external cues
SCT	2. Personal and cognitive factors	Factors relating to the individual and to forms of knowing and awareness
SCT	3. Environment	The external agents or conditions (physical, biological, social and cultural) that influence the behaviours of an organism
SCT	4. Symbolising capability	The capacity to use symbols to guide one's behavioural responses
SCT	5. Forethought capability	The ability to regulate behaviour based on future expectations
SCT	6. Vicarious capability	The ability to learn through observation of outcomes derived from the behaviours of others
SCT	7. Self-regulatory capability	The ability to regulate one's own behaviour on the basis of embedded skills and abilities
SCT	8. Self-reflective capability	The ability to analyse one's own experiences, thoughts and knowledge based on recalled occurrences
SCT	9. Self-efficacy	Perceived judgements of one's ability to cope effectively in different circumstances
SCT	10. Outcome expectations	Beliefs about the consequences of undertaking or not undertaking a behaviour based on their positive and negative value
TPB	1. Behaviour	The manifest, observable response in a given situation with respect to a given target
TPB	2. Intention	The cognitive representation of one's readiness to perform the behaviour; considered to be the immediate antecedent of behaviour
TPB	3. Attitude to the behaviour	Favourable or unfavourable evaluation or appraisal of the behaviour
TPB	4. Subjective norms	The perceived social pressure to perform or not to perform the behaviour
TPB	5. Perceived behavioural control	The perceived ability to perform the behaviour
TPB	6. Actual behavioural control	The actual ability to perform the behaviour
TPB	7. Behavioural belief composite	Sum of behavioural beliefs relating to a given behaviour
TPB	8. Behavioural belief[2]	Evaluation of a given outcome from a behaviour formed from the behavioural belief strength and outcome evaluation
TPB	9. Behavioural belief strength	The perceived probability that performing the behaviour will lead to the outcome in question
TPB	10. Outcome evaluation	Evaluation of a given outcome
TPB	11. Normative belief composite	Sum of normative beliefs
TPB	12. Normative belief[2]	Normative belief strength multiplied by motivation to comply
TPB	13. Normative belief strength	The perceived probability that the referent in question approves of the person performing the behaviour
TPB	14. Motivation to comply	The degree to which someone cares what a given social referent expects of them

TPB	15. Control belief composite	Sum of control beliefs
TPB	16. Control belief^2	Multiple of control belief strength and power of control factor
TPB	17. Power of control factor	The perceived ability of a given factor to facilitate or inhibit performance of the behaviour
TPB	18. Control belief strength	The perceived probability that the factor in question is present
TTM	1. Stage of Change	Stage of motivation/readiness to change behaviour
TTM	2. Precontemplation	Stage at which there is no serious consideration of behaviour change in the foreseeable future and during which one may be unaware of any need to / aware but unwilling to / defensive or resistant to / lack confidence in ability to change
TTM	3. Contemplation	Stage at which there is awareness of a problem and at which one is seriously considering behaviour change within the next six months but is not yet committed to act
TTM	4. Preparation	Stage at which one is ready for action and seriously intending to change within the next month
TTM	5. Action	Stage at which there is significant effort to change behaviour and where one has met a behaviour-specific criterion
TTM	6. Maintenance	Stage at which behaviour change has been sustained for more than six months and there is work to prevent relapse
TTM	7. Precontemplation-contemplation transition	Process of changing stage from Precontemplation to Contemplation
TTM	8. Contemplation-preparation transition	Process of changing stage from Contemplation to Preparation
TTM	9. Preparation-Action transition	Process of changing stage from Preparation to Action
TTM	10. Action-Maintenance transition	Process of changing stage from Action to Maintenance
TTM	11. Consciousness raising	The process of increasing awareness about the problem and improving the accuracy of information processing about the problem and about the self
TTM	12. Dramatic relief/emotional arousal	The process of experiencing and releasing feelings about the problem and the solution
TTM	13. Self re-evaluation	The process of cognitively and affectively assessing one's self-image in relation to the problem behaviour
TTM	14. Environmental re-evaluation	The process of cognitively and affectively assessing the ways in which a personal behaviour might have an impact on the social environment
TTM	15. Social liberation	The process of noticing social, policy or environmental changes that facilitate health behaviour change
TTM	16. Self liberation	The process through which one comes to believe in one's ability to change a particular behaviour and one's commitment to act on that belief
TTM	17. Stimulus control	The process through which the environment is restructured (e.g. by the individual) such that cues for problem behaviours are reduced and cues for healthier behaviours increased
TTM	18. Helping relationships	Relationships characterised by openness, trust and empathy, which are supportive with regard to the problem behaviour and health behaviour change
TTM	19. Counter conditioning	The process of adopting healthier behaviours as substitutes for problem behaviours
TTM	20. Reinforcement management	The process of rewarding oneself, or being rewarded by others, for making changes; contingency contracts, overt and covert reinforcement, self-reward
TTM	21. Decisional balance	The process through which the pros and cons of behaviour change are evaluated

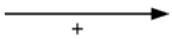
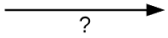
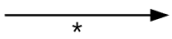
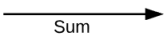
TTM	22. Self-efficacy	One's beliefs about one's ability to carry out a behaviour in any given situation
TTM	23. Temptation	A desire, or stimulus that creates a desire, to carry out the problem behaviour(s)

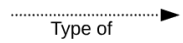
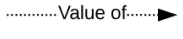
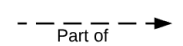
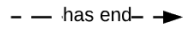
[1]HBM=Health Belief Model; SCT=Social Cognitive Theory; IMB=Information Motivation Belief model; TPB=Theory of Planned Behaviour; TTM=Transtheoretical Model. [2]The terms 'information' and 'attitude' were used by the theory authors but are intended to refer to information and attitudes to HIV-related issues.

Theory representations

Table 2 shows the accumulated library of relationships and their definitions. These were added incrementally as required to capture the theories, and are only a small set of possible relationships given that theories can specify relationships at any desired level of specificity.

Table 2: Relationships proposed in five theories of behaviour change

Type of Relationship	Label	Graphical Representation	Definition
Causal	Is influenced by		X is influenced by Y if, other things being equal, a change in Y necessitates a change in X
	Is influenced (+) by		X is influenced (+) by Y if, other things being equal, a change in the value of Y necessitates a change in the value of X in the same direction
	Is influenced (-) by		X is influenced (-) by Y if, other things being equal, a change in the value of Y necessitates a change in the value of X in the opposite direction
	May be influenced by		X may be influenced by Y if a change in Y necessitates a change in X only under some circumstances. The 'may be' can be combined with other modifiers such as (+) and (-) to express the fact that these more specific relationships hold true only under at least some circumstances
	Is influenced (*) by		X is influenced (*) by Y if X is a multiple of Y
	Is influenced (sum) by		X is influenced (sum) by Y if X is equal to the sum of values of Y where Y is a set with one or more members
	Correlates with		X is correlated with Y is values of X tend to covary with values of Y even though the two may

			not be causally linked. If X is correlated with Y, Y is correlated with X.
Semantic	Type of		X is a type of Y if X is a more specific subtype of Y, i.e. X possesses all the attributes of Y but Y does not possess all the attributes of X
	Value of		X is a value of Y if X is a type of Y, all types of Y are specified in the model and any use of a subtype of Y in the theory must be one of these
	Has attribute		X has attribute Y if Y is an attribute or characteristic of X
Structural	Part of		X is a part of Y if X is a member of a group whose members together comprise Y
	Has start		X has start Y if X is a transition and Y is a value at the temporal beginning of the transition
	Has end		X has end Y if X is a transition and Y is a value at the temporal end of the transition

Figures 1 to 5 show the five theories in terms of their constructs and the relationships between them. Tables 3 to 7 show the propositions in these theories. Each theory representation takes the form of a numbered list of theory propositions in the 'subject-relationship-object' format followed by a diagram. We have devised a number of conventions to interpret this diagram, which are shown in Table 3.

Health Belief Model (HBM)

Figure 1 and Table 3 show the Health Belief Model (HBM) represented using the proposed formal system. The HBM includes a 'moderating' relationship, whereby cues to action increase the influence of perceived threat on likelihood of preventive action. In the graphical representation, this is depicted by having the arrow ending on the perceived threat-likelihood of preventive action arrow. To represent this using a subject-relationship-object statement, we have to create a construct called the 'perceived threat-likelihood of preventive action relationship'. (See proposition 5 of Table 4.)

Figure 1 about here

Table 3: Propositions in the Health Belief Model

1. Likelihood of preventive action *is influenced (+) by* perceived benefits
2. Likelihood of preventive action *is influenced (-) by* perceived barriers
3. Likelihood of preventive action *is influenced (+) by* self-efficacy
4. Likelihood of preventive action *is influenced (+) by* perceived threat
5. Perceived threat-likelihood of preventive action relationship *is influenced (+) by* cues to action
6. Perceived benefits *are influenced by* demographic variables
7. Perceived benefits *are influenced by* social psychological variables
8. Perceived benefits *are influenced by* structural variables

9. Perceived barriers *are influenced by* demographic variables
10. Perceived barriers *are influenced by* social psychological variables
11. Perceived barriers *are influenced by* structural variables
12. Self-efficacy *is influenced by* demographic variables
13. Self-efficacy *is influenced by* social psychological variables
14. Self-efficacy *is influenced by* structural variables
15. Perceived severity *is part of* perceived threat
16. Perceived susceptibility *is part of* perceived threat

Information-Motivation-Behavioural Skills Model (IMB)

Figure 2 and Table 4 show the Information-Motivation-Behavioural Skills Model (IMB). We focused on the original version of the model which concerned HIV/AIDS prevention behaviour. The model has since been expanded to cover other health-risk behaviours [19].

The IMB includes the causal 'is influenced (*) by' relationship, and semantic 'type of' relationship. The latter expresses the idea that one construct is a subclass of another and logically inherits all its properties (e.g., mammal is a subclass of animal, and so mammals can be assumed to have all the properties of animals). Thus, the theory proposes that behavioural skills influence AIDS preventive behaviour and so it follows that all the different types of behavioural skill (e.g., self-efficacy) do so. Note that this relationship is different from the 'part of' relationship. For example, attitudes are deemed to be a part of motivation and so may contribute to it but not everything that is true about motivation would be true of attitudes.

The IMB also includes another semantic relationship: 'has attribute'. Thus, 'complexity' is represented as an attribute of AIDS preventive behaviour. This is used to express the proposition that the more complex the AIDS preventive behaviour is, the less information and motivation can influence it directly and the more it requires behavioural skills.

The IMB requires explicit representation of the idea that an entity may *or may not* influence another as a function of one or more stated or unstated moderators. We represent this by adding a ? to any given relationship. For example, it is an explicit feature of the theory that 'information' may or may not influence 'motivation'; therefore, we have represented 'Motivation *may be influenced by* information' by a solid arrow with a ?. Similarly, 'Enactment of AIDS preventive behaviour *may be influenced (+) by* information' is represented by a solid arrow with +?. In this case, the moderator 'complexity of AIDS preventive behaviour' is explicitly stated. From a logical standpoint this is not necessarily required, but from an accuracy perspective, it captures the original intent of the authors who proposed each theory.

In the graphical representation of the IMB we introduced a 'container' (a box around constructs). The container indicates that all contained constructs have the same relationship with another construct. This is a presentational device to make interpretation of the diagram easier.

Figure 2 about here

Table 4: Propositions in the Information-Motivation-Behavioural Skills Model

1. Enactment of AIDS preventive behaviour *is influenced (+) by* performance of behavioural skills
2. Enactment of AIDS preventive behaviour *has attribute* complexity of AIDS preventive behaviour
3. Enactment of AIDS preventive behaviour *may be influenced (+) by* information
4. Enactment of AIDS preventive behaviour *may be influenced (+) by* motivation

5. Information-enactment of AIDS preventive behaviour relationship *is influenced (+) by* complexity of AIDS related behaviour
6. Motivation-enactment of AIDS preventive behaviour relationship *is influenced (+) by* complexity of AIDS related behaviour
7. Performance of Behavioural skills *are influenced (+) by* information
8. Performance of Behavioural skills *are influenced (+) by* motivation
9. Motivation *may be influenced by* information
10. Information *may be influenced by* motivation
11. Self-acceptance of sexuality *is a type of* performance of behavioural skill
12. Ability to acquire behaviourally relevant information *is a type of* performance of behavioural skill
13. Self-efficacy *is a type of* performance of behavioural skill
14. Ability to practise self and partner reinforcement *is a type of* performance of behavioural skill
15. Ability to negotiate AIDS prevention with partner *is a type of* performance of behavioural skill
16. Ability to perform public prevention acts is *a type of* performance of behavioural skill
17. Consistent AIDS prevention ability *is a type of* performance of behavioural skill
18. Group-specific skills *are types of* behavioural skill
19. Information about means of transmission *are types of* information
20. Information about means of prevention *are types of* information
21. Decision rules about transmission and prevention *are types of* information
22. Implicit theories about risk and prevention *are types of* information
23. Attitudes *are part of* motivation
24. Social norms *are part of* motivation
25. Attitudes *are influenced by (*)* beliefs about outcomes
26. Attitudes are *influenced by (*)* evaluation of outcomes
27. Social norms *are influenced by (*)* beliefs about support/opposition from significant others
28. Social norms *are influenced by (*)* motivation to comply with referent others' wishes

Social Cognitive Theory (SCT)

Figure 3 and Table 5 show Social Cognitive Theory (SCT). The theory focuses on personal and cognitive factors that are important influences on behaviour and delineates these, expressing complex causal, structural and semantic relationships among them. The formalism, based on the theory propositions found, shows that according to those propositions self-efficacy is not explicitly considered a type of 'personal and cognitive factor' even though the definitions of these constructs suggests that it might be. This illustrates the ability of the formalism to make explicit aspects of theories that theory authors may have left implicit.

Figure 3 about here

Table 5: Propositions of Social Cognitive Theory

1. Behaviour *is influenced by* environment
2. Behaviour *is influenced (+) by* self-efficacy
3. Behaviour *is influenced by* personal and cognitive factors
4. Environment *is influenced by* behaviour
5. Environment *is influenced by* personal and cognitive factors
6. Personal and cognitive factors *are influenced by* environment
7. Personal and cognitive factors *are influenced by* behaviour
8. Self-efficacy *is part of* self-reflective capability
9. Self-efficacy *is influenced (+) by* self-regulatory capacity
10. Self-reflective capability *is a type of* personal cognitive factor

11. Self-regulatory capability *is a type of* personal cognitive factor
12. Self-regulatory capability *is influenced (+) by* vicarious learning capability
13. Self-regulatory capability *is influenced (+) by* outcome expectations
14. Symbolising capability *is a type of* personal cognitive factor
15. Forethought capability *is a type of* personal cognitive factor
16. Forethought capability *is influenced (+) by* symbolising capability
17. Vicarious learning capability *is a type of* personal cognitive factor
18. Outcome expectations *are part of* forethought capability

The Theory of Planned Behaviour (TPB)

Figure 4 and Table 6 show the Theory of Planned Behaviour (TPB). The TPB uses another type of causal relationship 'is influenced (sum) by' to represent the summation of multiple input values to influence the value of an output construct. Thus the construct 'behavioural belief composite' is the sum of a set of individual behavioural beliefs.

Each individual 'behavioural belief' is a multiple of the constructs 'behavioural belief strength' and 'outcome evaluation'. The theory author indicated in his comments that he did not perceive the need for the 'behavioural belief' construct representing each individual behavioural belief, but logically such a construct is required, otherwise there is nothing for the 'behavioural belief strength' and 'outcome evaluation' to multiply to create. We used a similar formulation for influences on 'normative belief composite' and 'control belief composite'.

It may be noted that 'perceived behavioural control' moderates 'intention-behaviour relationship'. This construct was created and labelled as part of the proposed formalism. The version of the theory presented here also includes the construct 'actual behavioural control' which captures the influence of capability and environmental factors on the' intention-behaviour relationship'.

One further addition required for this model was a 'correlates with' relationship. In our initial interpretation of the theory we construed this as a bi-directional causal relationship but the theory author indicated that this was not accurate. We therefore have to presume that 'correlates with' is a bi-directional relationship that proposes an association between two constructs that need not involve a direct cause. We represent this as a double-headed arrow. The TPB does not specify whether the correlations are positive or negative, but as with 'is influenced by' other theories might specify + or – qualifiers.

Figure 4 about here

Table 6: Propositions the Theory of Planned Behaviour

1. Behaviour *is influenced (+) by* intention
2. Intention *is influenced (+) by* attitude to the behaviour
3. Intention *is influenced (+) by* subjective norm
4. Intention *is influenced (+) by* perceived behavioural control
5. Intention-behaviour relationship *is influenced (*) by* perceived behavioural control
6. Intention-behaviour relationship *is influenced (*) by* actual behavioural control
7. Attitude to the behaviour *is influenced (+) by* summed behavioural beliefs
8. Behavioural belief composite *is influenced (sum) by* behavioural beliefs
9. Behavioural beliefs *are influenced (*) by* behavioural belief strengths
10. Behavioural beliefs *are influenced (*) by* outcome evaluations
11. Subjective norms *are influenced (*)* by normative belief composite
12. Normative belief composite *is influenced by (sum)* normative beliefs
13. Normative beliefs *are influenced (*) by* normative belief strength

14. Normative beliefs *are influenced (*)* by motivation to comply
15. Perceived behavioural control *is influenced (*) by* summed control beliefs
16. Perceived behavioural control *is influenced (+) by* actual behavioural control
17. Control belief composite *is influenced (sum) by* control beliefs
18. Control beliefs *are influenced (*) by* control belief strength
19. Control beliefs *are influenced (*) by* power of control factor
20. Attitude to the behaviour *correlates with* subjective norms
21. Attitude to the behaviour *correlates with* perceived behavioural control
22. Subjective norms *correlate with* perceived behavioural control

The Transtheoretical Model (TTM)

The final theory examined, the Transtheoretical Model (TTM) is different from the previous four in that it focuses on transitions between qualitatively different stages over time and factors that influence stage transitions (Figure 5 and Table 7). A person's stage of change is a measured variable, with each individual stage of change representing one of five possible values of the variable 'stage of change'. All constructs and relationships in the TTM end up influencing the value of the 'stage' variable. To represent this we used the 'value of' semantic relationship as a subclass of 'type of'. Without including 'value of' we would not be able to represent the proposition that different things influence different stage transitions.

The TTM also involves representing the transition from one stage to another as a construct. The transition constructs have 'has start' and 'has end' structural relationships. For example, our specification of the TTM includes the construct 'pre-contemplation to contemplation transition', which starts with 'pre-contemplation' and ends with 'contemplation'. For visual clarity, the graphical representation uses different shapes for constructs that are values and transitions.

Figure 5 about here

Table 7: Propositions of the Transtheoretical Model (TTM)

1. Pre-contemplation *is a value of* stage of change
2. Contemplation *is a value of* stage of change
3. Preparation *is a value of* stage of change
4. Action *is a value of* stage of change
5. Maintenance *is a value of* stage of change
6. Precontemplation contemplation transition *has start* precontemplation
7. Precontemplation contemplation transition *has end* contemplation
8. Contemplation preparation transition *has start* contemplation
9. Contemplation preparation transition *has end* preparation
10. Preparation action transition *has start* preparation
11. Preparation action transition *has end* action
12. Action maintenance transition *has start* action
13. Action maintenance transition *has end* maintenance
14. Precontemplation contemplation transition *is influenced (+) by* decisional balance
15. Precontemplation contemplation transition *is influenced (+) by* self-efficacy
16. Precontemplation contemplation transition *is influenced by* consciousness raising
17. Precontemplation contemplation transition *is influenced by* dramatic relief/emotional arousal
18. Precontemplation contemplation transition *is influenced by* self re-evaluation
19. Precontemplation contemplation transition *is influenced by* environmental re-evaluation
20. Precontemplation contemplation transition *is influenced by* social liberation
21. Contemplation preparation transition *is influenced (+) by* decisional balance
22. Contemplation preparation transition *is influenced (+) by* self-efficacy

23. Contemplation preparation transition *is influenced by* self re-evaluation
24. Contemplation preparation transition *is influenced by* environmental re-evaluation
25. Contemplation preparation transition *is influenced (+) by* self-liberation
26. Preparation action transition *is influenced (+) by* decisional balance
27. Preparation action transition *is influenced (+) by* self-efficacy
28. Preparation action transition *is influenced by* self-efficacy temptation combination
29. Preparation action transition *is influenced by* self-liberation
30. Preparation action transition *is influenced by* stimulus control
31. Preparation action transition *is influenced by* helping relationships
32. Preparation action transition *is influenced by* counter conditioning
33. Action maintenance transition *is influenced (+) by* decisional balance
34. Action maintenance transition *is influenced (+) by* self-efficacy
35. Action maintenance transition *is influenced by* self-efficacy temptation combination
36. Action maintenance transition *is influenced by* stimulus control
37. Action maintenance transition *is influenced by* helping relationships
38. Action maintenance transition *is influenced by* counter conditioning
39. Action maintenance transition *is influenced by* reinforcement management
40. Action maintenance transition *is influenced by* self-liberation
41. Self-efficacy temptation combination *is influenced by* self-efficacy
42. Self-efficacy temptation combination *is influenced by* temptation

Evaluation of theory specification

In all five theories, the theory authors appeared to be satisfied that the modelling system was able to capture their theories adequately. As noted earlier, the Theory of Planned Behaviour author did not consider that constructs such as 'behavioural beliefs' were required but logic dictated that these be specified and given a label since they captured a defined constructs within the theory. The Transtheoretical Model author noted that the model as represented here did not capture transition back to earlier stages of change and processes that might influence these but noted that this would be possible using the formal system we were using.

Discussion

Our formal system was found to be capable of representing the five selected behaviour change theories. The process of deriving the representations helped to clarify interpretations of the theories through dialogue with theory authors and experts.

The theory representation method chosen was one of many possible options. Some behaviour change theories are efficiently expressed as sets of equations, while others are naturally expressed in terms of micro- or macro-simulations such as Agent Based Models, Bayesian nets or Markov processes [20]. We believe that the chosen system can encompass those other systems as well as theories that want to retain the expressiveness of natural language.

Another outcome from the process of specifying theories is highlighting the need for greater clarity and consistency in defining constructs. The definitions obtained from the literature involved a great deal of implicit meaning and in some cases did not appear to capture fully the construct as intended. It is also apparent that some constructs are common to more than one theory, although whether they are precisely the same or whether the theory authors consider there to be important differences is not clear.

Developing this theory representation system highlighted the importance in theories of being more precise about the nature of complex causal interactions. The Theory of Planned Behaviour, for

example, contains two different kinds of causal interaction: 1) 'multiplicative' relationships between factors contributing to behavioural beliefs, subjective norms and perceived behavioural control, and 2) a 'moderating' relationship of 'actual behavioural control' over the 'intention-behaviour relationship'. Both of these can be handled in this modelling system and in principle any complexity of interaction can be handled. For example, in principle someone could specify another variable moderating the moderating relationship between 'actual behavioural control' and the 'intention-behaviour relationship'. In practice, such higher-order interactions are very challenging to test in empirical studies but the potential to represent them in this formalism exists.

Once a database of theories represented using the proposed system is created, computational methods can be used for searching, comparing and integrating. Possible operations include:

1. Simple searching for constructs that have similar labels or definitions: for example to find all theories that use the construct label, 'self-efficacy' or 'intention'.
2. 'Semantic searching' to find constructs that have similar properties even though they may have different labels and definitions.
3. Searching for properties of constructs: for example what are the factors that have been proposed to influence self-efficacy and what has it been postulated to influence?
4. Precisely delineating areas of overlap and difference between theories: for example, specifying precisely how recently developed theories such as I-Change [21] compare to earlier ones such as the Theory of Planned behaviour and the Health Belief Model.
5. Clustering theories to identify theory classes: for example, based on similarity indices based on overlap in constructs and proposed relationships.
6. Building a 'canonical' theory that includes key relationships and constructs from the corpus of theories: for example, by creating an aggregated list of constructs, selecting ones that are most representative of sets that are similar, and building a specification that captures all their proposed inter-relationships.
7. Reasoning with theories to assess differential predictions from different theories covering the same domain: for example, specifically predicting which constructs would be expected to impact on behaviour change in defined behaviours, populations and settings.

The current study was limited in terms of the number of theories that it covered. It is possible that there are theories in the literature that could not be specified using this system. However, the theories included to date were diverse and we are already advanced with work to specify another 77 theories which we will report on in a subsequent publication.

Another potential limitation is that the formal system has not yet been formally linked to modelling languages used in other domains such as OWL (Web Ontology Language; www.w3.org/OWL/). These modelling languages allow potentially greater expressivity than the system used here and include a large set of formalisms and axioms that can be used to generate automated inferences. This is something that will done in parallel with the next phase of the project.

A third limitation is that the theory representations were only used to capture the main parts of the theories. We believe that it is straightforward to add further propositions to the theories but the graphical representations would likely become very difficult to read. Finding a way to present complex theories in graphical form is always going to be challenging but merits serious consideration, including input from specialists across different disciplines.

Conclusions

It appears to be possible to represent major behaviour change theories using a standard system of binary relationships between defined constructs. This system could be used both for reformulating existing theories and for developing new ones. This would set the scene for a more systematic approach to theory development and use. The implication for scientific advance is that the system allows for comparison across the large number of behaviour change theories to identify the frequently occurring constructs and relationships, enabling the development of a smaller number of canonical theories and providing a coherent basis for advance. It allows for more specific testing of theories and reporting of findings in a fashion that supports the efficient accumulation of knowledge. An important implication for practice is that identifying a few theories that summarise core theoretical knowledge about behaviour change will enable those wishing to apply theory, but are currently overwhelmed with the large number of theories, to use those summary theories to develop and evaluate interventions and to inform syntheses of evidence about behaviour change.

Method

Theory selection

We selected five theories from the 83 reviewed by Davis et al [3,14]. They were the most frequently used behaviour change theories as identified by previous reviews of behaviour change interventions for a range of health behaviours, including healthy eating, physical activity, and sexual behaviours [3,15,22]: the Health Belief Model (HBM) [23], the Information-Motivation-Behavioural Skill Model (IMB) [24], Social Cognitive Theory (SCT) [25], the Theory of Planned Behaviour (TPB) [26], and the Transtheoretical Model (TTM) [27].

Choice of a formal representation system

Behaviour change theories describe causal influences between variables such that a change in one variable results in a change in another. They also describe processes of change in which events lead to other events or values of a variable change over time as a function of other variables and events. These causal influences may themselves be influenced or 'moderated' by other variables or processes, and variables or processes may 'mediate' causal influences between other variables or processes.

Models of this kind can be represented as sets of expressions of the form 'subject-relationship-object'. The subject and object are the constructs, and the relationship links the subject to the object. Subject-relationship-object expressions can be represented graphically by labelled boxes with labelled arrows linking them. They can be used to create formal 'ontologies': data structures that represent knowledge in terms of 'entity' identifiers, labels, definitions and relationships between them [28].

Figure 6 gives a simple example of a graphical representation of a simple behaviour change theory represented in this way. The arrows represent the relationship 'positively influences'. The diagram represents dyadic relationships between variables for a given behaviour: habit strength positively influences behaviour frequency, desire strength positively influences behaviour frequency, and opportunity frequency positively influences behaviour frequency.

Figure 6 about here

Causal models represented in this way are used in many areas of study including psychology, economics, biology, and public health [29]. They allow direct and indirect paths of influence to be traced between variables. The relationships in subject-relationship-object expressions need not be

limited to causal ones; they can represent any relationship in which two constructs may be linked with each other (e.g., 'is a type of' as in 'disgust is a type of emotion').

The ability to form expressions linking two constructs with any type of relationship can provide a flexible and intuitive system which could be suitable for representing behaviour change theories. We applied it to the five selected behaviour change theories using a two-step process. For each theory: 1) we identified labels and definitions for each construct in each theory, and 2) we expressed propositions within the theory in terms of defined relationships between the constructs, building a library of defined relationships as required.

A database of theories constructed in this way can provide the basis for search, comparison, integration, development and use. The database can be searched for construct labels or definitions; similarities and differences can be identified; relationships with other constructs can be compared; propositions can be combined or merged; specific propositions can be extracted from theories and tested; propositions can be added, removed or modified; and causal chains can be used to identify intervention targets.

Construct identification and definition

We identified construct labels (e.g., 'self-efficacy') from the original published descriptions of the theories [14]. Construct definitions were generated in several steps. They were drafted by KS, RC and/or LC (see authors and acknowledgements) based on the natural language expressions of the theories. They were then reviewed by two additional authors. If no definition was found within the theoretical statements, the dictionary of the American Psychological Association [30] was consulted. If no definition was found here, other dictionaries and sources were consulted (e.g., Oxford English Dictionary, Merriam-Webster Dictionary, Wikipedia). If it was not possible to find an appropriate definition using any of these sources, definitions were created drawing upon the original theory text and dictionary definitions for words within the construct label.

The draft labels and definitions were reviewed by the study leads, RW and SM, and then sent to the theory authors to review for accuracy. For authors with whom contact could not be made, theory experts were identified and contacted. These experts were found by searching the Web of Science and Scopus databases for authors who had cited the theory most frequently, and then examining published articles by these authors for relevance. Where theory authors or experts proposed modification, addition or removal of theory constructs, these changes were either included or an explanation provided as to why not.

Theory representation

To express propositions within the theories, we identified three broad types of relationship that could be used: causal, semantic, and structural. A *causal* relationship would be of the kind 'X is influenced by Y'. If X and Y are variables this means that if values of X change, then other things being equal, so do values of Y. If X is a process it means that X leads to Y. The nature of the dependency may be more or less specific and may take different forms as in 'X is positively influenced by Y' or 'X is proportional to Y'. What we are calling a *semantic* relationship would be of the kind 'X is a type of Y' or 'X has attribute Y'. A *structural* relationship would be of the kind 'X is a part of Y'. Table 8 gives examples of the three types of relationship.

Table 8: Basic types of relationship

Type of relationship	**Basic relationship**	**Examples**

Causal	influenced by	*Biology:* Height is influenced by level of production of growth hormone *Behaviour change:* Strength of intention to perform a behaviour is influenced by attitude towards the behaviour
Semantic	type of	*Biology:* Dogs are a type of mammal *Behaviour change:* Information about consequences of a behaviour is a type of information
Structural	part of	*Biology:* Ribs are a part of the skeleton *Behaviour change:* Perceived severity is a part of perceived threat

Each type of relationship is directional. Some could, in principle, be expressed in the reverse direction (e.g., 'influences' versus 'is influenced by'). The choice of direction in those cases is determined by how one wants the model to operate. In the case of behaviour change models the aim is generally to predict and explain behaviour and so we judged it to be natural to make the entity being explained or predicted as the subject of each causal relationship (variants of 'is influenced by' rather than 'influences').

To represent the constructs and relationships comprising a theory in an accessible visual format, we created a graphical system of representation using Lucid Chart software (www.lucidchart.com). A useful feature of this software is that the resulting diagrams can be exported to a database as a list of 'subject-relationship-object' triads (e.g., 'behaviour frequency', 'is positively influenced by', 'strength of desire').

Two researchers (SM and CG) independently created formal representations of the five theories in terms of subject-relationship-object triads. Each theory was then discussed with the other members of the research team until an agreed version was arrived at. This diagram (with accompanying theory representation in the form of subject-relationship-object propositions) was then sent to theory authors or experts for comment and revised to reflect their views.

Data availability

N/a

Code availability

N/a

Figure captions

Figure 1: Example of a simple causal model of behaviour

Figure 2: Representation of the Health Belief Model

Figure 3: Representation of the Information Motivation Behavioural Skills Model (IMB)

Figure 4: Representation of Social Cognitive Theory

Figure 5: Representation the Theory of Planned Behaviour

Figure 6: Representation of the Transtheoretical Model (TTM)

References

1 Green, L. W., Hiatt, R. A. & Hoeft, K. S. in *Oxford Textbook of Global Public Health* (eds R Detels, M Gulliford, Q Karim, & C Chuan) 218-233 (Oxford University Press, 2015).

2 Vlek, C. & Steg, L. Human Behavior and Environmental Sustainability: Problems, Driving Forces, and Research Topics. *Journal of Social Issues* **63**, 1-19 (2007).

3 Davis, R., Campbell, R., Hildon, Z., Hobbs, L. & Michie, S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health psychology review* **9**, 323-344 (2015).

4 Glantz, K., Rimer, B. & K, V. (Jossey-Bass, California, 2015).

5 Craig, P. *et al.* Developing and evaluating complex interventions: the new Medical Research Council guidance. *British Medical Journal* **337**, doi:10.1136/bmj.a1655 (2008).

6 Glanz, K. & Bishop, D. B. The role of behavioral science theory in development and implementation of public health interventions. *Annual review of public health* **31**, 399-418 (2010).

7 Lippke, S. & Ziegelmann, J. P. Theory-based health behavior change: Developing, testing, and applying theories for evidence-based interventions. *Applied Psychology* **57**, 698-716 (2008).

8 West, R. & Michie, S. *A Guide to the Development and Evaluation of Digital Behaviour Change Interventions*. (Silverback Publishing, 2016).

9 Armitage, C. J. & Conner, M. Social cognition models and health behaviour: A structured review. *Psychology and health* **15**, 173-189 (2000).

10 Gebhardt, W. A. & Maes, S. Integrating social-psychological frameworks for health behavior research. *American Journal of Health Behavior* **25**, 528-536 (2001).

11 Kirscht, J. P. Preventive health behavior: A review of research and issues. *Health Psychology* **2**, 277 (1983).

12 Michie, S. *et al.* Making psychological theory useful for implementing evidence based practice: a consensus approach. *BMJ Quality & Safety* **14**, 26-33 (2005).

13 Garnett, C. *et al.* Reported Theory Use by Digital Interventions for Hazardous and Harmful Alcohol Consumption, and Association With Effectiveness: Meta-Regression. *Journal of medical Internet research* **20**, e69, doi:10.2196/jmir.8807 (2018).

14 Michie, S., West, R., Campbell, R., Brown, J. & Gainforth, H. *ABC of Behaviour Change Theories: An Essential Resource for Researchers, Policy Makers and Practi*. (Silverback Publishing, 2014).

15 Prestwich, A. *et al.* Does theory influence the effectiveness of health behavior interventions? Meta-analysis. *Health Psychology* **33**, 465-474 (2014).

16 Dombrowski, S. U., Sniehotta, F. F., Avenell, A. & Coyne, J. C. Towards a cumulative science of behaviour change: do current conduct and reporting of behavioural interventions fall short of best practice. *Psychol Health* **22**, 869-874 (2007).

17 Michie, S. & Prestwich, A. Are interventions theory-based? Development of a theory coding scheme. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association* **29**, 1-8, doi:10.1037/a0016939 (2010).

18 Cane, J., O'Connor, D. & Michie, S. Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implementation science : IS* **7**, 37, doi:10.1186/1748-5908-7-37 (2012).

19 Fisher, W., Fisher, J. & Shuper, S. *Advances in Experimental Social Psychology*. (2014).

20 El-Sayed, A. & Galea, S. (Oxford University Press, Oxford, 2017).

21 Vries, H. D. & Mudde, A. N. Predicting stage transitions for smoking cessation applying the attitude-social influence-efficacy model. *Psychology and Health* **13**, 369-385 (1998).

22 Painter, J. E., Borba, C. P., Hynes, M., Mays, D. & Glanz, K. The use of theory in health behavior research from 2000 to 2005: a systematic review. *Annals of Behavioral Medicine* **35**, 358 (2008).

23 Rosenstock, I. M. Historical origins of the health belief model. *Health education monographs* **2**, 328-335 (1974).
24 Fisher, J. D. & Fisher, W. A. Changing AIDS-risk behavior. *Psychological bulletin* **111**, 455 (1992).
25 Bandura, A. Social foundation of thought and action: A social-cognitive view. *Englewood Cliffs* (1986).
26 Ajzen, I. The theory of planned behaviour. Organizational Behaviour and Human Decision Processes, 50, 179-211. *De Young*, 509-526 (1991).
27 DiClemente, C. C. *et al.* The process of smoking cessation: an analysis of precontemplation, contemplation, and preparation stages of change. *Journal of consulting and clinical psychology* **59**, 295 (1991).
28 Hastings, J. in *The Gene Ontology Handbook* 3-13 (Springer, 2017).
29 Judea, P. *Causality: Models, Reasoning, and Inference*. (Cambridge University Press, 2000).
30 VandenBos, G. R. *APA dictionary of psychology*. (American Psychological Association, 2007).

Funding and acknowledgements
We are particularly grateful to the theory authors and experts who commented on the construct definitions and theory representations: Icek Ajzen (TPB), James Annesi (SCT), Carlo DiClemente (TTM), William Fisher (IBM), and Vic Strecher (HBM). We are grateful to Kate Sheals (KS) for assistance with coding the theories and drafting the paper. We are also grateful to Marie Johnston for early discussions about this piece of work and to Marta Marques, Alison Wright, Claire Garnett, Artur Direito, Hilary Groarke, Holly Walton, and Tracey Revenson who have kindly assisted with the drafting of the paper. This work was also supported by a grant from the UK Medical Research Council (grant number MR/L011115/1) and by a National Institute of Health Research Senior Investigator's award held by SM. RW's contribution was part funded by Cancer Research UK. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests
RW has undertaken research and consultancy from companies that develop and manufacture smoking cessation medicines (Pfizer, GSK, J&J).

Contributions
RW contributed to study conception, revised coding of theories and led on preparation of the manuscript. SM conceived the project, and contributed to manuscript preparation. LCB, RC and CG undertook initial coding of theories, contributed to study conception and commented on manuscript drafts. CG, JH, and CL contributed to manuscript preparation.