Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data)

Susan Michie,[1]* Caroline E Wood,[1] Marie Johnston,[1,2] Charles Abraham,[3] Jill J Francis[4] and Wendy Hardeman[5]

[1]Research Department of Clinical, Educational and Health Psychology, University College London, London, UK
[2]Institute of Applied Health Sciences, College of Life Science and Medicine, University of Aberdeen, Aberdeen, UK
[3]University of Exeter Medical School, University of Exeter, St Luke's Campus, Exeter, UK
[4]School of Health Sciences, City University London, London, UK
[5]Primary Care Unit, Cambridge Institute of Public Health, University of Cambridge, Cambridge, UK

*Corresponding author

Declared competing interests of authors: Charles Abraham acknowledges funding from the UK National Institute for Health Research and the Peninsula Collaboration for Leadership in Applied Health Research and Care.

Published November 2015
DOI: 10.3310/hta19990

Scientific summary

Behaviour change techniques

Health Technology Assessment 2015; Vol. 19: No. 99
DOI: 10.3310/hta19990

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Interventions to change behaviour are 'co-ordinated sets of activities designed to change specified behaviour patterns' (Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011;**6**:42). They are often complex and involve multiple, interacting components. Lack of a methodology for unpacking this 'black box' prevents progress in developing more effective interventions and in reliably implementing effective ones. Detailed specification of interventions is important for basic and applied behavioural science. A methodology for achieving this is to specify intervention content in terms of their 'active ingredients', that is, behaviour change techniques (BCTs). The UK Medical Research Council (MRC) guidance for developing and evaluating complex interventions advocates this approach (Craig P, Dieppe O, MacIntyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;**337**:a1655).

Structured lists, otherwise known as 'taxonomies', of BCTs have been developed to enable the clear specification of intervention content in terms of BCTs. The first cross-domain BCT taxonomy consisted of 22 BCTs and four BCT packages. Taxonomies have been developed for different behavioural domains: physical activity, healthy eating, smoking, alcohol consumption and condom use. Using taxonomies with standardised labels, clear definitions and examples will improve current practice in that any given BCT will be described by the same label and that label will be used to describe the same BCT. There is a need for consensus about labels and definitions that can be understood across disciplines, behavioural domains and countries. A key objective of the programme of research reported here was to develop such a consensus and to facilitate the use of a shared language to describe BCTs.

Objectives

The aim was to develop a method for specifying interventions to change behaviour in terms of component BCTs. Objectives were as follows.

Phase 1: development

Develop a taxonomy made up of BCTs that:

i. are applicable to an extensive range of behaviour change interventions (BCIs)
ii. are agreed by an international consensus to be potential active components of BCIs
iii. are clearly labelled and well defined to promote reliable application by researchers and intervention developers from a range of disciplines and countries
iv. are distinct (non-overlapping, non-redundant) and precise, and can be used with confidence to describe interventions
v. are hierarchically organised to improve ease of use.

Phase 2: evaluation

i. Assess the effectiveness of two programmes of taxonomy user training in improving reliable identification of BCTs.
ii. Assess and report the reliability and validity of using the developed taxonomy to code intervention descriptions into BCTs.
iii. Assess whether or not use of the taxonomy leads to clearer and more replicable written intervention descriptions.

Phase 3: supporting use of the developed taxonomy (Behaviour Change Technique Taxonomy version 1)

Develop accessible resources to facilitate reliable, valid and confident use of the developed taxonomy including:

i. a website providing regular updates on progress of the Behaviour Change Technique Taxonomy version 1 (BCTTv1) project
ii. an online, interactive training resource for users of BCTTv1
iii. a smartphone application version of BCTTv1 available across mobile platforms.

Cross-phase: implementation and dissemination

Ensure that development of the taxonomy:

i. is supported by a range of multidisciplinary and international researchers, practitioners, funders, journal editors and policymakers, and
ii. lays foundations for co-ordinated advances in behavioural science.

Empirical studies included in the project

Phase 1: development

Study 1: developing a comprehensive list of behaviour change techniques

Study 1 developed an extensive, consensually agreed list of BCTs used in BCIs. First, a list of distinct BCT labels and definitions was systematically developed using Delphi methods, building on a preliminary list integrating six published taxonomies (Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. *Health Psychol* 2008;**27**:379–87; Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol* 2009;**28**:690–701; Michie S, Hyder N, Walia A, West R. Development of a taxonomy of behavior change techniques used in individual behavioral support for smoking cessation. *Addict Behav* 2011;**36**:315–19; Michie S, Whittington C, Hamoudi Z, Zarnani F, Tober G, West R. Identification of behavior change techniques to reduce excessive alcohol consumption. *Addiction* 2012;**107**:1431–40; Dixon D, Johnston M. *Health Behavior Change Competency Framework: Competences to Deliver Interventions to Change Lifestyle Behaviors that Affect Health*. Edinburgh: Scottish Government; 2012; Abraham C. Mapping Change Mechanisms and Behaviour Change Techniques: A Systematic Approach to Promoting Behaviour Change Through Text. In Abraham C and Kools M (editors) *Writing Health Communication: An Evidence-Based Guide for Professionals*. London: SAGE Publications; 2012). The list was refined following feedback from the study's multidisciplinary International Advisory Board (IAB) of 30 behaviour change experts. BCTs were added, divided and removed, and labels and definitions refined to capture the smallest components compatible with retaining the proposed active ingredients with the minimum of overlap. This resulted in a taxonomy of 93 distinct BCTs: BCTTv1.

© Queen's Printer and Controller of HMSO 2015. This work was produced by Michie *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Study 2: Behaviour Change Technique Taxonomy version 1 with hierarchical structure

The 93-item BCT list poses problems for easy recall of, and ready access to, the BCTs and thus its speed and accuracy of use. Study 2 aimed to provide a structure for the list to increase ease of use and to compare a pragmatic 'bottom-up' method with a theoretical 'top-down' method for generating structure. The objectives were to (1) develop a hierarchical structure using an open-sort task ('bottom-up' method), (2) identify whether or not BCTs could reliably be linked to domains of theoretical constructs using a closed-sort task ('top-down' method) and (3) identify any overlap between 'bottom-up' and 'top-down' groupings. The 'bottom-up' method resulted in 16 groupings, which had some overlap with the theory-driven groupings.

Phase 2: evaluation

Study 3: Behaviour Change Technique Taxonomy version 1 training development and evaluation

The process of coding interventions into BCTs is a highly skilled task requiring familiarity with labels and definitions and ability to make a series of complex interpretative judgements. This requires an effective programme of coder training to ensure sufficient level of skills to demonstrate reliability and validity of identifying BCTs. Reliability can be measured by the extent to which coders agree with each other on the presence and absence of BCTs in interventions while validity can be measured by the extent to which coders agree with expert consensus (i.e. the judgement consensus of the BCTTv1 study team). Study 3 developed two training programmes: 1-day workshops and distance group tutorials. Effectiveness was evaluated in terms of whether or not training enhanced coding reliability, validity, confidence in identifying BCTs and coding competence. Both methods of training improved trainee agreement with expert consensus, doubling the proportion achieving a satisfactory level of performance and confidence in identifying the BCTs assessed but did not improve intercoder agreement. Training was evaluated positively by trainees with both types of training receiving uniformly high ratings in terms of usefulness.

Study 4: identifying behaviour change technique in interventions using Behaviour Change Technique Taxonomy version 1 – reliability and validity

Study 4 investigated the reliability and validity of BCTTv1 when used by trained coders (from study 3). Intercoder and test–retest reliability (measured at two time points, 1 month apart) were used to assess reliability of using BCTTv1 to code BCTs in published intervention descriptions. Validity was assessed by the extent to which coding agreed with consensus by the BCTTv1 study team as to which BCTs were present. We also investigated trained coders' confidence in identifying BCTs and whether or not descriptions varied in the ease with which BCTs could be identified. Good intercoder reliability was observed for 80 of the 93 BCTs. Reliability did not differ across the two time points and there was good test–retest reliability. Of the 15 BCTs agreed by expert consensus to be present in the descriptions, good agreement was achieved for 14, thus indicating good validity.

Study 5: reporting intervention descriptions using Behaviour Change Technique Taxonomy version 1 – clarity and distinctiveness

The final study, consisting of three substudies, evaluated the extent to which using BCTTv1 improved clarity of reporting descriptions that had been observed and compared this in trained and untrained users. Writers were shown videos of BCIs as delivered in practice and wrote descriptions. Descriptions generated were then rated for quality (i.e. clarity, ease of understanding and ease of replicability), recognition of BCTs and the extent to which different interventions could be recognised and distinguished. Results were mixed: providing BCTTv1 plus training resulted in descriptions of poorer quality when comparing trained with untrained writers (study 5b) but resulted in improvement when investigated in a within-person study (study 5c), while only providing the taxonomy made no difference (study 5a). However, reliability of BCT identification was greater for descriptions written by trained than untrained writers, with BCTTv1 (study 5b) but not in the within-person study (study 5c). Providing BCTTv1 both with and without training did not improve validity of BCT identification.

Design

The five empirical studies used the following methods:

- Delphi methods with feedback from a multidisciplinary IAB to generate a list of distinct BCT labels and definitions.
- Open-sort methods to categorise BCTs and generate a hierarchical structure.
- Comparison of user's intercoder reliability using prevalence- and bias-adjusted kappa (PABAK; for which PABAK > 0.60 is considered good reliability), agreement with expert consensus (in this case, 'expert' is defined as individuals with at least 15 years' experience of researching and coding interventions) and user confidence in identifying BCTs after, compared with before, being trained.
- Assessment of agreement between trained coders (intercoder reliability), agreement over time (test–retest reliability) and trained coder agreement with expert consensus (validity).
- Within- and between-participant experimental designs to compare quality of intervention descriptions reported by untrained and trained writers, with and without access to BCTTv1.

Participants

A total of 400 participants were involved across the studies (note: some participants took part in more than one study).

Study 1 ($n = 41$)

Nineteen international behaviour change experts, active in their field and engaged in investigating, designing and/or delivering interventions, 16 of the 30 members of the BCTTv1 project IAB, five BCTTv1 team members and a 'lay' person.

Study 2 ($n = 36$)

Eighteen of the 19 international experts from the first study and 18 additional experts with experience of designing interventions, writing manuals or protocols and/or conducting narrative or systematic reviews of behaviour change literature.

Study 3 ($n = 161$)

One-hundred and sixty one trainee coders (systematic reviewers, researchers, practitioners and policy-makers engaged in investigating, reviewing, designing and delivering interventions).

Study 4 ($n = 40$)

Forty trained coders who had also taken part in the third study.

Studies 5 (a, b, c) ($n = 190$)

One hundred and sixty-six trainee intervention reporters (systematic reviewers, practitioners, policy-makers engaged in investigating, reviewing, designing, delivering or describing interventions), 12 smoking cessation practitioners with no experience of using BCT taxonomies, 12 trained coders from studies 3 and 4.

© Queen's Printer and Controller of HMSO 2015. This work was produced by Michie *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Results

Study 1

The BCTTv1 is the first cross-behaviour taxonomy to be established by an international group of experienced BCT coders. It comprises 93 distinct, clearly labelled, well-defined BCTs hierarchically organised into 16 groupings.

Study 2

Groupings were created by those with experience of BCIs: average of 15.11 groupings (standard deviation 6.11, range 5–24 groupings). BCTs relating to 'Reward and Punishment' and 'Cues and Cue Responses' were perceived as markedly different from other BCTs. Fifty-nine of the BCTs were reliably allocated to 12 of the 14 theoretical domains; 47 were significant and 12 were of borderline significance. There was a significant association between the 16 'bottom-up' groupings and the 13 'top-down' groupings, chi-squared = 437.80; $p < 0.001$. Thirty-six of the 208 'bottom-up' × 'top–down' pairings (i.e. 16 × 13) showed greater overlap than expected by chance. However, only six combinations achieved satisfactory evidence of similarity. The moderate overlap between the groupings indicates some tendency to implicitly conceptualise BCTs in terms of the same theoretical domains.

Study 3

Training was effective in improving application of BCTTv1. Trainee agreement with expert consensus improved across workshop and distance group tutorial trainees (both $p < 0.05$), with the proportion achieving competence (i.e. agreement with expert consensus PABAK > 0.60) doubling (both $p < 0.05$). Forty-six per cent of workshop trainees and 78% of tutorial trainees achieved competence after training. Trainee confidence in identifying BCTs assessed improved in workshops ($p < 0.001$). Intercoder agreement did not improve for either method (workshops $p = 0.08$ and tutorials $p = 0.57$).

Study 4

The BCTTv1 can be used accurately and reliably by trained coders to specify BCTs in BCIs: good intercoder reliability was observed across 80 BCTs identified in the protocols: 64 of the 80 BCTs (i.e. 80%) achieved a mean PABAK score of > 0.70 and 59 of the 80 BCTs (i.e. 74%) achieved a mean PABAK score of > 0.70. Good test–retest reliability was achieved within coders (between baseline and 1 month) ($p < 0.001$). Reliability was poorer for frequent BCTs; 9 out of the 16 frequently identified BCTs failed to achieve a good level of reliability, that is a PABAK score of > 0.70. Good concurrent validity was achieved for 14 out of the 15 BCTs identified by experienced coders to be present in published descriptions.

Studies 5 a, b, c

It is unclear whether or not BCTTv1 benefits reporting of descriptions of observed interventions.

Study 5a: quality, reliability of BCT identification (between trained coders) and validity (trained coder agreement with expert consensus) did not differ for descriptions written by untrained writers with and without BCTTv1 (all $p > 0.05$).

Study 5b: quality of descriptions written by untrained raters without BCTTv1 was rated as greater than descriptions written by trained writers, with BCTTv1 (all $p < 0.05$). Reliability of BCT identification (between trained coders) was greater for descriptions written by trained writers, with BCTTv1 (mean PABAK trained = 0.87; untrained = 0.84; $p < 0.05$).

Study 5c: quality of descriptions written by trained writers, with BCTTv1, was higher than descriptions written by the same writers, before training and not using BCTTv1 (all $p < 0.05$ except replicability of mode of delivery). There were no differences in reliability ($p = 0.50$) or validity ($p = 0.28$).

Conclusions

With regard to BCTTv1, 'v1' refers to version 1, to indicate that we envisage development of this taxonomy in the future. It is an extensive, hierarchically organised taxonomy of 93 distinct BCTs with clear definitions and examples that offer a generally reliable method for specifying the active content of interventions. Eighty out of these 93 were found in 40 published intervention descriptions, indicating the need for such an extensive list of BCTs. Both methods of BCTTv1 training improve reliability and validity for most BCTs. BCTTv1 can be used by trained coders to identify BCTs accurately in intervention descriptions. However, there is room for improvement in terms of clarifying some BCT definitions and in developing further training to improve reliability and validity. It is not yet clear whether or not BCTTv1 enhances reporting as our evaluation method only used observed interventions. BCTTv1 represents a useable method for research and practitioner communities when interpreting and implementing BCIs.

Funding

This project was funded by the MRC Ref: G0901474/1. Funding also came from the Peninsula Collaboration for Leadership in Applied Health Research and Care.

© Queen's Printer and Controller of HMSO 2015. This work was produced by Michie *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Health Technology Assessment

HTA MRP–MRC

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.116

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme or, commissioned/managed through the Methodology research programme (MRP), and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

This issue of the Health Technology Assessment journal series contains a project commissioned/managed by the Methodology research programme (MRP). The Medical Research Council (MRC) is working with NIHR to deliver the single joint health strategy and the MRP was launched in 2008 as part of the delivery model. MRC is lead funding partner for MRP and part of this programme is the joint MRC–NIHR funding panel 'The Methodology Research Programme Panel'.

To strengthen the evidence base for health research, the MRP oversees and implements the evolving strategy for high-quality methodological research. In addition to the MRC and NIHR funding partners, the MRP takes into account the needs of other stakeholders including the devolved administrations, industry R&D, and regulatory/advisory agencies and other public bodies. The MRP funds investigator-led and needs-led research proposals from across the UK. In addition to the standard MRC and RCUK terms and conditions, projects commissioned/managed by the MRP are expected to provide a detailed report on the research findings and may publish the findings in the HTA journal, if supported by NIHR funds.

The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded under a MRC–NIHR partnership. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the MRC, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the MRC, NETSCC, the HTA programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2015. This work was produced by Michie *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Editor-in-Chief of *Health Technology Assessment* and NIHR Journals Library

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson Director of NETSCC, HTA, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Professor Elaine McColl Director, Newcastle Clinical Trials Unit, Institute of Health and Society, Newcastle University, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Health Sciences Research, Health and Wellbeing Research and Development Group, University of Winchester, UK

Professor John Norrie Health Services Research Unit, University of Aberdeen, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Institute of Child Health, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk