Enhancing Behavior Change Technique Coding Methods: Identifying Behavioral Targets and Delivery Styles in Smoking Cessation Trials

Nicola Black, PhD1,6 • A. Jess Williams, MSc1,2 • Neza Javornik, MSc1 • Claire Scott, MSc1,3 • Marie Johnston, PhD1 • Maarten C. Eisma, PhD1,4 • Susan Michie, DPhil5 • Jamie Hartmann-Boyce, DPhil6 • Robert West, PhD7 • Wolfgang Viechtbauer, PhD8 • Marijn de Bruin, PhD1

Abstract

Background The behavior change technique (BCT) taxonomy v1 is often used in systematic reviews for identifying active components of interventions. Its utility could be enhanced by linking BCTs to specific target behaviors and qualifying BCT delivery style.

Purpose To determine whether behavioral targets and delivery styles of BCTs can be coded reliably and to determine the utility of coding these characteristics.

Methods As part of a large systematic review of 142 smoking cessation trials, two researchers independently coded publicly and privately held intervention and comparator group materials, specifying the behavioral target (quitting, abstinence, medication adherence, or treatment engagement) and delivery style (tailored vs. not tailored; active participation vs. passive receipt) of each BCT.

Results Researchers coded 3,843 BCTs, which were reliably attributed to behavioral targets (AC1 = 0.92, PABAK = 0.91). Tailoring (AC1 = 0.80, PABAK = 0.74) and participation (AC1 = 0.71, PABAK = 0.64) were also coded reliably. There was considerable variability between groups in quitting and abstinence BCTs (ranges: 0–41; 0–18) and in tailoring and participation (ranges: 0–20; 0–32), but less variability for medication adherence and treatment engagement (ranges: 0–6; 0–7).

Conclusions Behavioral targets and delivery styles of BCTs can be reliably identified and occur with sufficient frequency in smoking cessation trials for inclusion in quantitative syntheses (e.g., meta-regression analyses). Systematic reviewers could consider adopting these methods to evaluate the impact of intervention components targeting different behaviors, as well as the benefits of different BCT delivery styles.

Keywords Behavior change technique • Smoking cessation • Reliability • Systematic review • Delivery style • Tailoring

Introduction

Health risk behaviors such as tobacco smoking are important causes of disease and disability [1], health care expenditure [2], and societal costs related to loss of labor [3]. Numerous behavioral interventions have been, and continue to be, developed, evaluated, and published; however, suboptimal reporting of these interventions limits their implementation, replication,
Many BCTs in the BCTTv1 are defined in such a way that they may or may not be tailored or require active participant engagement. For example, BCT 5.1 “information about health consequences” may include personalized information based on the participant’s assessed health status (tailored), general health consequences (not tailored), be delivered via collaborative discussion (active), or be delivered via a leaflet (passive). These delivery styles were included in an earlier BCT taxonomy for medication adherence [18, 19] but have not yet been applied to BCTs used in smoking cessation trials. Identifying these two styles of BCT delivery (tailoring and participation) is an important step in fully specifying the active content of smoking cessation interventions.

Enhancing investigation of BCT effects by considering BCT coding in relation to coding of behavioral targets and style of BCT delivery could benefit both evidence syntheses and intervention development and delivery. Those synthesizing evidence on interventions could not only identify the associations between BCTs and intervention effectiveness, but also identify which BCTs are effective in promoting behavior initiation versus behavior maintenance, whether intervention components targeting auxiliary behaviors such as medication adherence and treatment engagement result in better outcomes, and whether the effectiveness of BCTs varies depending on the delivery style. Those developing and delivering interventions could use these systematic reviews to gain a clearer picture of the content of effective interventions than would be offered by systematic reviews that only specify the BCTs used. This is likely to increase the replicability of effective interventions and their active components.

There is general consensus that it is important to achieve adequate inter-coder reliability on BCTs extracted from published intervention descriptions, and this has been demonstrated for the majority of BCTs included in the BCTTv1 [20]. In this study, we will examine whether (i) BCTs can be reliably attributed to specific target behaviors and whether the BCT delivery styles can be reliably identified, and (ii) these behaviors and delivery styles are occurring with sufficient frequency to make this additional data extraction work useful for enhancing intervention replicability and meta-analyses.

### Method

#### Design

This study is part of a larger, ongoing review of smoking cessation trials (IC-SMOKE; PROSPERO registration number CRD42015025251 [21]). Full details, including the data, of all outputs from the IC-SMOKE project will be available on the project’s Open Science Framework page (https://osf.io/23hfv/) upon publication.

The Cochrane Tobacco Addiction Group Specialized Register was searched for randomized controlled trials...
(RCTs) assessing the impact of behavioral interventions (with or without pharmacological support) on biochemically verified smoking cessation at 6 months or longer. Trials were excluded if they were published before 1996 or after November 1, 2015 (the search date), were not reported in English or in peer-reviewed journals, or if the participants were aged under 18 years [21].

**Data Extraction**

Data were first collected from the publicly available materials (e.g., the primary trial articles and appendices, but also protocols and additional publications such as intervention development papers). In addition, a comprehensive procedure was used for contacting authors of all included trials to request additional materials describing their experimental and comparator interventions to obtain privately held materials (e.g., intervention manuals, practitioner training materials, websites, self-help materials [21]). While not pertinent to the current study as it does not allow for calculating reliability or identifying tailoring and participation, it should be noted that authors were also asked to complete a brief checklist detailing the active content delivered to the comparator group in their trial (discussed further in Discussion). Publicly available materials—besides the primary article, which was retrieved for all groups—were retrieved for 59% of groups (intervention: 61%; comparator: 56%) and privately held materials were retrieved for 45% of groups (intervention: 51%; comparator: 37%). When checklist responses are included, privately held materials/information were retrieved for 64% of groups (intervention: 63%; comparator: 65%). The procedure for retrieving additional materials took approximately 8 months, with one final response received at 11 months.

Identifying BCTs, their behavioral targets, and delivery styles in these materials involved two steps: independent coding and discussion to resolve disagreements. First, two independent researchers coded materials for the presence of each of the 93 BCTs in the BCTTv1 [4]. Minor adaptations to the taxonomy were made prior to coding, including the removal of BCT 8.5 (overcorrection), as it was unclear how this would be used in this context, the addition of BCT 4.5 (tell to act, defined as “tell the person to perform the target behavior”), and the inclusion of smoking cessation examples for each BCT, to enhance validity and inter-coder reliability. For efficiency, the BCT coding sheet was designed such that the additional properties could be coded without additional note taking (see definitions of these properties in Table 1 and a simplified illustration of the coding sheet in Fig. 1). Researchers read the materials, and each time a new BCT was identified, the BCT code and source quote were entered into the first two columns. It was possible for a single quote to contain multiple BCTs. This was followed by columns for tailoring and participation for BCTs targeting each behavior. Researchers entered a T or A into the tailoring and participation columns of the targeted behavior if the BCT was tailored or actively delivered, respectively. Smoking cessation interventions often define a formal quit date. To manage workload, coders were instructed to code a BCT no more than once for each behavior before the quit date and no more than once after the quit date (even if it occurred more frequently). Hence, a total possible 93 (BCTs) × 4 (behaviors) × 2 (before/after quit date) = 744 BCTs could be coded per (intervention or comparator) group. Researchers were also asked to identify any BCTs that were particularly difficult to code, to inform potential sensitivity analyses. In the second step, researchers used discussion to resolve discrepancies in BCT codes. The BCT coding took approximately 10 weeks, including discussion time.

**Data Analysis**

To examine whether the behavioral targets and delivery style of BCTs could be reliably identified, two indicators of inter-rater reliability were used: Gwet’s AC1 [22] and

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Elements of the enhanced behavior change technique coding scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavior</strong></td>
<td><strong>Coding</strong></td>
</tr>
<tr>
<td>Quitting</td>
<td>BCTs used to increase the likelihood of the participant ceasing tobacco smoking (initiating a quit attempt)</td>
</tr>
<tr>
<td>Abstinence</td>
<td>BCTs used to increase the likelihood of the participant maintaining their non-(tobacco-)smoker state</td>
</tr>
<tr>
<td>Medication adherence</td>
<td>BCTs used to increase the likelihood of the participant using their smoking cessation medication in appropriate dosages at appropriate times</td>
</tr>
<tr>
<td>Treatment engagement</td>
<td>BCTs used to increase the likelihood of the participant engaging with, and completing components of, the smoking cessation treatment</td>
</tr>
<tr>
<td><strong>Tailoring</strong></td>
<td><strong>Participation</strong></td>
</tr>
<tr>
<td>Tailored</td>
<td>The BCT was modified based on characteristics of the recipient</td>
</tr>
<tr>
<td>Active</td>
<td>The delivery of the BCT required the participant to actively participate</td>
</tr>
</tbody>
</table>

*BCT* behavior change technique.
prevalence- and bias-adjusted kappa (PABAK [23]). AC1 and PABAK were chosen because they are more stable indicators of inter-rater reliability than is the widely used Cohen’s kappa [24, 25]. Results were interpreted using Altman’s guidelines: ≤0.20 poor, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 good, and 0.81–1.00 very good reliability [26]. These analyses were conducted on BCTs independently identified by both coders prior to resolving discrepancies in BCT coding. Analyses were conducted separately for each group type (intervention or comparator) as well as overall. This was to pick up any potential differences in coding ability of the content provided for the intervention and comparator groups.

To evaluate the utility of coding behavioral targets and delivery style of BCTs, we examined whether these properties occurred with sufficient frequency for inclusion in quantitative meta-regression analyses. Utility was judged according to whether researchers would be able to examine questions such as “Does using BCTs to target behavior X (i.e., cessation, abstinence, adherence, engagement) improve outcomes (e.g., smoking cessation rates, quit attempts, medication adherence, attrition)?” “Does tailoring intervention content improve outcomes?” and “Does active client participation improve outcomes?” To answer these questions, histograms were inspected to determine whether there was sufficient spread across groups in (a) the number of BCTs used to target each of quitting, abstinence, treatment engagement, and, amongst those groups who received medication, medication adherence, (b) the number of BCTs that were tailored, and (c) the number of BCTs that involved active client engagement. Variables with little spread would not be useful as predictor variables in meta-regression analyses. Histograms are presented separately by group type (intervention or comparator), as it was expected that intervention groups would tend to contain more BCTs, more tailored BCTs, and more actively delivered BCTs. Because only about half of the studies reported a quit date, for consistency between trials, quit date was removed from the dataset for this analysis, thus allowing each group to provide one BCT per target behavior (for a possible total of 93 BCTs × 4 behaviors = 372 BCTs per group).

Results

Descriptives

One hundred forty-two studies reporting 204 intervention and 142 comparator groups were included. Included studies are listed in the Supplementary Material. Through coding the publicly available and privately held materials, there were 3843 BCTs (intervention: 2860; control: 983) that were identified by both of the researchers, 725 BCTs were identified by one of the researchers but not the other (388/725 identified by researcher one but not two; 337/725 identified by researcher two but not one; this constitutes 0.3% [725/(744 BCTs × 346 groups)] of the total number of judgments that were made by each researcher), and 4128 BCTs after discrepancies were reconciled, and quit date was removed from the dataset (for reasons mentioned earlier). The 3843 BCTs were used for the reliability analyses and the 4128 BCTs for the utility analyses.

Table 2 presents the mean number of BCTs agreed present by both researchers, by treatment arm and targeted behavior. An average of 14.82 BCTs per intervention group and 7.78 per comparator group were coded.

Reliability of Linking BCTs to Behaviors and Delivery Style

As seen in Table 3a, reliability of attributing BCTs to one of four behaviors was very good (0.91–0.94). Reliability was good to very good for identifying tailoring of BCTs (0.73–0.84) and good for identifying participation in the delivery of BCTs (0.64–0.74).
During BCT coding, the two researchers identified particular difficulties with identifying the delivery style of one BCT: BCT 3.1 Social support (unspecified). When this BCT was removed (n = 420 instances) from the reliability analyses, reliability for coding tailoring (0.73–0.88) and participation (0.75–0.85) improved, whereas reliability of coding the behavioral targets remained the same (0.91–0.94; see Table 3b).

Utility of Coding Targeted Behaviors and Delivery Style of BCTs

The degree of variability between different intervention and comparator groups in the use of BCTs is displayed in Table 2 and the histograms in Figs. 2 and 3. As seen in Fig. 2, between intervention groups, there was considerable variability in BCTs targeting quitting and abstinence, but minimal variability in BCTs targeting medication adherence and treatment engagement. Between comparator groups, there was considerable variability in BCTs targeting quitting, but minimal variability in BCTs targeting abstinence, medication adherence, and treatment engagement. As seen in Fig. 3, there was considerable variability in tailored and active BCTs between intervention groups, but more limited variability between comparator groups.

Discussion

This study examined the reliability and utility of a BCT coding scheme that extends beyond extracting exclusively the presence or absence of a BCT in intervention descriptions. For a sample of 142 smoking cessation trials (346 intervention and comparator groups), we examined published materials and additional materials obtained from study authors. Behavioral targets and BCT delivery style could be identified with good to very good reliability. The utility of extracting these data for use in meta-regression analyses was evident for quitting and abstinence in relation to the target behaviors, and for tailoring and participation in relation to delivery styles, but less so for medication adherence and treatment engagement. Hence, this study demonstrated that extending BCT coding to include specific behavioral targets and styles of BCT delivery is feasible and adds substantial information to the coding of BCT occurrence only, which is currently the most common practice.

It appears useful to collect information on the style of BCT delivery and on BCTs targeting quitting and abstinence delivered to intervention groups, but less so for BCTs targeting medication adherence and treatment engagement. It might be that these techniques are infrequently used, infrequently reported, or both. If they are used but infrequently reported, this limits the ability to replicate published interventions and to synthesize evidence on effective intervention techniques. If they are infrequently used, this highlights two areas of trial and intervention development in need of improvement. Fewer than half (46%) of the intervention groups who received medication received any behavioral support to help them adhere to its intended use. Given that medication adherence remains a challenge [27], trial developers in the field of smoking cessation should consider the use of appropriate BCTs to promote adherence. Similarly, low treatment engagement

### Table 2 Mean (SD) number of BCTs coded by targeted behavior and group

<table>
<thead>
<tr>
<th></th>
<th>Quitting</th>
<th>Abstinence</th>
<th>Medication adherence</th>
<th>Treatment engagement</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>11.28 (8.35)</td>
<td>2.40 (3.84)</td>
<td>0.70 (1.26)</td>
<td>0.44 (1.18)</td>
<td>14.82 (10.85)</td>
</tr>
<tr>
<td>Comparator</td>
<td>6.61 (7.60)</td>
<td>0.64 (1.72)</td>
<td>0.41 (0.95)</td>
<td>0.12 (0.61)</td>
<td>7.78 (8.93)</td>
</tr>
</tbody>
</table>

BCT behavior change technique.

### Table 3 Inter-rater reliability for coding targeted behavior, tailoring, and participation of behavior change techniques

<table>
<thead>
<tr>
<th></th>
<th>Intervention</th>
<th>Comparator</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AC₁</td>
<td>PABAK</td>
<td>AC₁</td>
</tr>
<tr>
<td>a. All BCTs (n = 3,843)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavior</td>
<td>0.92</td>
<td>0.91</td>
<td>0.94</td>
</tr>
<tr>
<td>Tailoring</td>
<td>0.78</td>
<td>0.73</td>
<td>0.84</td>
</tr>
<tr>
<td>Participation</td>
<td>0.69</td>
<td>0.64</td>
<td>0.74</td>
</tr>
<tr>
<td>b. All BCTs except 3.1 (n = 3,423)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavior</td>
<td>0.92</td>
<td>0.91</td>
<td>0.94</td>
</tr>
<tr>
<td>Tailoring</td>
<td>0.80</td>
<td>0.73</td>
<td>0.88</td>
</tr>
<tr>
<td>Participation</td>
<td>0.80</td>
<td>0.75</td>
<td>0.85</td>
</tr>
</tbody>
</table>

BCT behavior change technique; PABAK prevalence- and bias-adjusted kappa.
leads to attrition, which can contribute to incomplete delivery of intervention content and biased estimates of intervention effectiveness [28]. Trial developers should thus also consider how BCTs could be used to retain participants in the intervention and trial. Alternatively, it might be that the use of BCTs to promote treatment engagement is reactive; treatment providers might utilize BCTs only when they observe that a participant is becoming disengaged. In this case, authors should capture such information in published trial reports.

Usual practice when conducting meta-regressions using BCTs has been to code the presence of each BCT targeting any of a cluster of final health behaviors (e.g., dietary behaviors), without identifying the delivery styles of each BCT, or disentangling which of the final health behaviors (e.g., vegetable intake, fat intake) or preparatory behaviors (e.g., buying food, preparing food) is being targeted [7, 8]. Current findings suggest that only limited additional information would be gained if reviewers were to widely adopt the coding of behavioral targets such as medication adherence and treatment engagement (though, this could be due to poor reporting in existing trials). Comparatively, coding the use of BCTs targeting behavior change (e.g., quitting) and behavior

Fig. 2 Histograms showing the degree of between-group variability in use of behavior change techniques, by group type and behavioral target.
change maintenance (e.g., abstinence) could be a useful addition to reviews of smoking cessation trials and reviews of interventions for other health behaviors, such as substance use, diet, and physical activity. Theoretical accounts of behavior would suggest that BCTs that shift the relative cost–benefit analysis in favor of the new behavior should promote behavior change initiation (e.g., promoting the perceived benefits and/or reducing the perceived costs should promote behavior change [11, 12]). Comparatively, behavior change maintenance could be promoted through BCTs that promote habit, resource availability and utilization, positive maintenance motives, supportive environments, and self-regulation to monitor behavior and overcome barriers [29].

The coding scheme presented in the current study would allow systematic reviewers to assess (a) whether authors use theoretically supported BCTs at each stage of behavior change and (b) whether these BCTs are effective.

Comparator groups received noticeably fewer BCTs than did intervention groups. This finding will partially reflect reality, in that intervention group support can be expected to be more intensive than comparator group support. It is plausible that this finding will also partially reflect differing reporting qualities for intervention and comparator group support. Comparator group support tends to be poorly reported and, in the case of usual care comparator groups, not manualized [18, 19]. For these reasons, we have also developed a checklist based on previous work [18, 19] for collecting information from authors on the support provided to comparator groups and will be reporting on the data collected using this method elsewhere. Such information was not relevant to the current study, as it does not allow for calculating the reliability of coding behavioral targets and delivery styles of BCTs, nor for the assessment of the delivery styles of BCTs. Nonetheless, it is worth briefly mentioning that the apparent number of BCTs delivered to comparator groups increases considerably when data from this checklist are included.

Strengths of this study are the rigorous systematic review methodology applied and the use of a considerable number of unpublished intervention materials that were obtained through contacting authors. Materials from 142 trials and 346 intervention and comparator groups were coded independently by two researchers, and bias-and prevalence-corrected reliability calculations were used given the skewed distributions. Limitations of this study are that BCT coding was conducted by two trained researchers and the degree to which other teams are able to reliably extract these data has yet to be examined. Further, a single behavioral domain (smoking cessation) was examined; it may be that extracting behaviors and delivery styles in other domains is more or less difficult. Nonetheless, our findings suggest that exploring whether different behavioral targets and delivery styles can be reliably and usefully
identified in other behavioral domains is warranted. Finally, only the first instance (before and after quit date) was coded for each BCT. This procedure is also likely to have resulted in an underestimation of the utility of the methods described, given that fewer BCTs were coded overall. Future researchers could avoid this problem by coding for repeated delivery of BCTs, when this occurs. Despite these limitations, it is important to note that much more BCT data were collected than is the case in most systematic reviews using the taxonomy. Further, to the authors’ knowledge, the data presented here provide the most comprehensive representation to date of the active content of behavioral smoking cessation interventions.

To our knowledge, this study represents the first attempt at moving the coding of BCTs delivered to intervention and comparator groups beyond presence or absence of BCTs. It presented and tested an enhanced coding scheme for characterizing this active content of behavioral interventions. The proposed extensions can be coded reliably, and, on the whole, these extensions are likely to be useful to both those attempting to replicate effective interventions and those trying to synthesize the evidence on behavior change interventions.

**Supplementary Material**

Supplementary material is available at *Annals of Behavioral Medicine* online.

**Acknowledgments** We thank the authors of included studies who responded to our requests for materials on the intervention and comparator group support. These materials were crucial to achieving our current aims. We also thank the members of our advisory board panels who provided valuable input into the broader study design. This work was funded by Cancer Research UK (application number C50862/A18446). The systematic review protocol was previously peer-reviewed by Cancer Research UK as part of the funding process. The funders had no role in protocol design, decision to publish, or preparation of the manuscript.

**Compliance with Ethical Standards**

**Authors’ Statement of Conflict of Interest and Adherence to Ethical Standards** Author R. West’s salary was funded by Cancer Research UK. N. Black’s, A.J. Williams’s, M.C. Eisma’s, and C. Scott’s salaries were funded by Cancer Research UK. RW undertakes research and consultancy for companies that develop and manufacture smoking cessation medications (Pfizer, J&J, and GSK). Advisory board: R. West and S. Michie are unpaid advisors to the UK’s National Centre for Smoking Cessation and Training. All other authors declare that they have no conflicts of interest.

**Authors’ Contributions**

The study was designed by M. de Bruin, M. Johnston, S. Michie, M. C. Eisma, J. Hartmann-Boyce, R. West, and W. Viechtbauer. Data were collected by N. Black, A. J. Williams, N. Javornik, M. C. Eisma, and C. Scott. The final analyses were conducted by N. Black, A. J. Williams, and N. Javornik. N. Black, A. J. Williams, M. de Bruin, and M. Johnston drafted the manuscript, which was critically revised by all authors. All authors approved the final manuscript.

**References**


