



HHS Public Access

Author manuscript

West J Nurs Res. Author manuscript; available in PMC 2022 August 09.

Published in final edited form as:

West J Nurs Res. 2021 May ; 43(5): 468–477. doi:10.1177/0193945920954782.

Applying the Behavior Change Technique Taxonomy to four multi-component childhood obesity interventions

Meghan M JaKa, PhD* [Evaluation Associate],

HealthPartners Institute, Center for Evaluation and Survey Research, Minneapolis, MN, US

Caroline Wood, PhD [Honorary Senior Researcher],

Centre for Digital Public Health in Emergencies, University College London, UK

Sara Veblen-Mortenson, MPH/MSW [Program Manager in the Dean's Office],

School of Public Health at the University of Minnesota, MN, US

Shirley M. Moore, PhD, RN [Edward J. and Louise Mellen Professor of Nursing Emerita; Distinguished University Professor Emerita],

Frances Payne Bolton School of Nursing, Case Western Reserve University, Cleveland, OH, US

Donna Matheson, PhD [Senior Research Scientist],

Department of General Pediatrics, Stanford Medical School, Stanford University, Stanford, CA, US

June Stevens, PhD [Distinguished Professor],

Departments of Nutrition and Epidemiology at the University of North Carolina, Chapel Hill, NC, US

Lou Atkins, PhD [Senior Teaching Fellow],

Centre for Behaviour Change, University of London, UK

Susan Michie, BA, MPhil, DPhil [Professor of Health Psychology and Director],

Centre for Behaviour Change at University College London, UK

Clara Adegbite-Adeniyi, BA [Research Project Manager],

Case Western Reserve University, Frances Payne Bolton School of Nursing, Cleveland, OH, US

Oluwatomisin Olayinka, PhD,RN, MPH [Assistant Professor],

MGH Institute of Health Professions School of Nursing, Boston, MA, US

Eli K. Po'e, BS [Research Assistant],

Vanderbilt University Medical Center, Nashville, TN, US

Alethea M. Kelly, BS [Budget Accounting Analyst Pre-Award],

Vanderbilt University Medical Center, Nashville, TN, US

Holly Nicastro, PhD, MPH [Program Director],

Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, Bethesda, MD, US

Shrikant I. Bangdiwala, PhD [Professor, Director of Statistics],

*Corresponding Author: Meghan M. JaKa, PhD is an Evaluation Associate at the HealthPartners Institute, Center for Evaluation and Survey Research, Minneapolis, MN, US with no potential conflicts of interest to declare. [Meghan.m.jaka@healthpartners.com](mailto: Meghan.m.jaka@healthpartners.com).

Department of Health Research Methods, Evidence and Impact, Population Health Research Institute, McMaster University, Hamilton, ON, CA

Shari L. Barkin, MD, MSHS [Professor of Pediatrics],
Vanderbilt University Medical Center, Nashville, TN, US

Charlotte Pratt, PhD, RD [Health Scientist Administrator and Deputy Branch Chief],
National Heart, Lung, and Blood Institutes, National Institutes of Health, Bethesda, MD, US

Thomas N. Robinson, MD, MPH [Irving Schulman, MD Endowed Professor in Child Health, Professor of Pediatrics and of Medicine, Director],
Center for Healthy Weight at Stanford University and Lucile Packard Children's Hospital Stanford, Stanford, CA, US

Nancy E. Sherwood, PhD [Associate]
Division of Epidemiology and Community Health, School of Public Health, Minneapolis, MN, US
with no potential conflicts of interest to declare.

Abstract

Applying the Behavior Change Technique Taxonomy has the potential to facilitate identification of effective childhood obesity intervention components. This article evaluates the feasibility of coding Childhood Obesity Prevention and Treatment Consortium interventions and compares reliability between external Taxonomy-familiar coders and internal intervention-familiar coders. After training, coder pairs independently coded pre-specified portions of intervention materials. An adjudication process was used to explore coding discrepancies. Reliability between internal and external coders was moderate (prevalence and bias-adjusted kappa .38 to .55). Reliability for specific target behaviors varied with substantial agreement for physical activity (.63 to .76) and moderate for dietary intake (.44 to .63). Applying the Taxonomy to these interventions was feasible, but agreement was modest. Coding discrepancies highlight the importance of refining coding to capture the complexities of childhood obesity interventions, which often engage multiple recipients (e.g., parents and/or children) and address multiple behaviors (e.g., diet, physical activity, screen time).

Keywords

childhood obesity; behavior change interventions; intervention design; intervention measurement; fidelity

Childhood obesity is a global public health problem (Ng et al., 2014). Unfortunately, childhood obesity prevention and treatment trials often yield null or modest results (Psaltopoulou et al., 2019). Obtaining a better understanding of the inner-workings and nuanced design of childhood obesity interventions could lead to the development of more effective interventions (Tate et al., 2016). The etiology of obesity is multifaceted and thus interventions necessarily target several weight-related behaviors, employ multiple behavior change techniques and target multiple participants (e.g., parents and children) concurrently. This complexity contributes to the challenge of identifying the components of effective interventions and understanding why interventions are ineffective. Combining results from

multiple intervention studies is also more difficult when their intervention components are poorly specified (Bangdiwala et al., 2016).

Tools to Specify Behavioral Interventions

To help researchers better specify intervention components, behavior change experts have created a number of common language hierarchies. These tools help identify commonalities in seemingly disparate interventions and ultimately uncover cross-study findings. For example, one intervention may be described as using ‘goal setting’ and another as using ‘intention planning’ when in fact both are deploying the same approach. Using different terminology to describe similar intervention techniques prevents readers from easily comparing across studies to understand the true impact of such intervention strategies. Via these common language tools, researchers can conduct cross-study reviews to identify which techniques are associated with outcomes and then design studies to explicitly test the inclusion or exclusion of various strategies to optimize interventions (Black et al., 2020; Martin et al., 2013).

One commonly used tool is the Behavior Change Technique Taxonomy, a hierarchically organized, cross-behavior taxonomy created to describe behavior change techniques, the active ingredients designed to bring about change in an intervention (Michie et al., 2013). The Taxonomy was developed using a rigorous Delphi method with numerous behavior change experts compiling techniques across many prior classification systems. The validity of this approach has been established and the reliability and utility have been demonstrated across multiple fields since its inception (Abraham et al., 2015; Michie et al., 2015; Wood et al., 2016). Coding techniques involve a deductive process of categorizing intervention content and making interpretative judgments about the presence or absence of each specific technique.

Recent research has shown that coders can gain competence in coding techniques by completing coder training (Abraham et al., 2015; Wood et al., 2015). There is a growing interest in applying the Taxonomy to childhood obesity interventions, but the majority of studies to date have not captured nuanced details about which techniques have been used with which target behavior (e.g., diet, physical activity) and to whom techniques have been delivered (i.e., parents alone, children alone, both parents and children) (Martin et al., 2013). Being able to specify the presence or absence of techniques delivered within target behaviors and delivered to target participants will help experts in the field examine the relative effectiveness of these techniques. Further, training study staff to identify techniques in their own intervention materials can help to reduce the cost of coding, increase specificity and consistency of intervention reporting across the field, and support the uptake of this rigorous methodology more broadly. However, it is unknown whether the coder’s degree of familiarity with a particular intervention and their degree of experience with the Taxonomy influences coding. Given this, the effectiveness of the online Taxonomy training program for teaching study-specific coders to code intervention materials should be evaluated.

The Childhood Obesity Prevention and Treatment Research Consortium

The Childhood Obesity Prevention and Treatment Research (COPTR) consortium provided a unique opportunity to address important questions regarding the application of the Taxonomy to childhood obesity interventions (Pratt et al., 2013). The consortium included four study centers testing four unique interventions designed with varying intervention content and some common design elements such as targeting multiple weight-related behaviors, inclusion of multiple components and settings, and inclusion of parents and children as target participants (Moore et al., 2013; Po'e et al., 2013; Robinson et al., 2013; Sherwood et al., 2013). Results reported to date indicate variability in the impact of the COPTR interventions on obesity and related health behaviors (Barkin et al., 2018; French et al., 2018; Moore et al., 2019). Thus, using the Taxonomy to characterize the interventions has the potential to shed light on behavior change techniques that are differentially effective.

Purpose

In the present study we evaluated inter-coder reliability between COPTR intervention team coders (internal coders) and coders from the Centre for Behaviour Change at University College London (external coders) for coding a year of intervention materials for each of the four unique COPTR interventions. The aims of this study were to: describe the degree of inter-coder agreement between internal coders and external coders; describe the degree of inter-coder agreement between internal and external coders across target behaviors (diet, physical activity, screen time, sleep, stress) and target participants (parent, child, both, or unspecified); and describe the degree of inter-coder agreement within internal coder pairs and within external coder pairs at specified intervals. An adjudication process was then utilized to understand the discrepancies between internal and external coders in order to inform recommendations for coding childhood obesity intervention materials.

Methods

To examine the inter-coder reliability between internal coder pairs and between external coder pairs, intervention materials for each of the COPTR interventions (e.g., manuals for intervention delivery, participant handouts, phone scripts) were coded using the Behavior Change Technique Taxonomy (Michie et al., 2013). Below are descriptions of the interventions and associated materials. The existing online Taxonomy training program and process are also described.

Minnesota's NET-Works Intervention

The NET-Works intervention promoted healthy eating and physical activity among preschool-aged children (Sherwood et al., 2013). The intervention was delivered in three components that included home visits, parenting classes, and pediatric primary care. Home visits were conducted monthly over three years and assisted parents in developing parenting skills and encouraged healthy eating, physical activity and reduced television viewing among children. Home visits included goal setting, activities to promote healthy skills and norms, and the promotion of linkages to related community resources. Home visitors also provided between-visit support by telephone or e-mail. Parenting classes, conducted by

parent educators, complemented home visits with a matched curriculum of topics such as parenting skills, healthy eating, active play and limitations on television viewing. Finally, home visitors worked in partnership with parents to promote conversations with primary care practitioners on key weight-related topics at annual well-child visits. The NET-Works intervention was associated with lower levels of total energy intake and television viewing time at 2 and 3-year follow-ups. In secondary analyses, NET-Works significantly reduced BMI (body mass index) over three years among Hispanic children and children with baseline overweight or obesity (French et al., 2018).

Case Western Reserve University's IMPACT Intervention

IMPACT was a three-arm trial including two intervention conditions and a control (Moore et al., 2013). The HealthyCHANGE intervention included in this analysis draws on motivational interviewing and cognitive-behavioral techniques to facilitate behavior change. Participants received information about diet, physical activity, sleep and stress management; small group intervention sessions; and phone contacts. The other intervention arm, not reviewed here, was the SystemsCHANGE intervention which focused on restructuring family environment and a series of family-designed small experiments. Neither of the family-based interventions improved BMI slope (primary outcome) over 3 years or health-related secondary outcomes (Moore et al., 2019).

Vanderbilt's GROW Healthier Intervention

GROW Healthier was a multi-level behavioral intervention directed at parent-preschooler pairs and their built environment (Po'e et al., 2013). It was a 3-year family-centered intervention focused on skills-building around healthy lifestyles with three phases: intensive, maintenance, and sustainability. The 3-month intensive phase was conducted at the community recreation center with 12 weekly skills-building sessions in nutrition, physical activity, and parenting. The 9-month maintenance phase included monthly coaching phone calls that reinforced skills taught in the intensive phase. The 2-year sustainability phase included cues to action to use the built environment to support healthy behaviors for parents and children. Participants received a BMI feedback graph annually and an adaptive intervention phone call if they had moved into an overweight or obese classification. Participation in the GROW Healthier intervention was associated with reductions in total energy intake but not BMI trajectory over the 3-year follow-up period (Barkin et al., 2018).

Stanford's GOALS Intervention

GOALS was a multi-level (individual children, families, groups, primary care providers, and community youth-serving organizations), multi-component (eating behaviors, including environmental changes in the home, physical activity/inactivity, and screen time), and multi-setting (homes, community centers, and primary care clinics) intervention. Intervention components included home visits, an afterschool team sports program, and primary care provider progress reports, and were delivered over three full years. Families chose the order they wished to address eating behaviors, physical activity or screen time and each family progressed at their own rate through content as skills were mastered. The intervention was grounded in social cognitive theory (Bandura, 1997) with emphasis on Latino cultural values (Resnicow et al., 1999). Research in cognitive and social psychology was used to

frame intervention components promoting greater and more sustained effects (Dweck, 2017; Robinson, 2010).

Behavior Change Technique Taxonomy Online Training

The online Taxonomy training (www.bct-taxonomy.com) is based on tutorial training and includes practice coding tasks, feedback, structured discussion led by expert tutors, access to support networks and additional resources. The training progressively introduces the 44 most common techniques using hands-on tasks with intervention descriptions from journal articles across various health behaviors. The training does not cover all 93 techniques, assuming coder ability to apply learnings to the remainder of less-common techniques. Six training sessions provide technique descriptions, a new coding guideline and an interactive task. Coders are given a score and automated feedback after each session. The final two sessions are only available after scoring 70% on the first assessment. Recommend training time is 1 hour per session and 1 session per week.

Coder Training, Certification, Background and Experience

Internal Coders—Internal coding pairs, two from each COPTR study, had at least an undergraduate degree in health-related fields. None had prior experience coding the Taxonomy. The internal coders had variable levels of familiarity with their study's intervention materials, including developing materials or providing intervention supervision, but none directly delivered interventions. All internal coders completed the online Taxonomy training and achieved at least 70% in the first assessment. As part of the training, internal coders completed the online training and participated in three 1-hour discussions led by expert Taxonomy tutors from University College London. Prior to the webinar, internal coders practiced coding using one of the COPTR intervention design manuscripts (Sherwood et al., 2013) as well as sample intervention materials from a past childhood obesity trial (Robinson et al., 2007).

External Coders—External coders were four Health Psychology post-graduate students at University College London (two pairs). All external coders had previously completed the online training and had experience using the Taxonomy to specify techniques in descriptions of health-related interventions. External coders did not participate in the three additional training discussions with Taxonomy tutors. For consistency, the pairs also coded the single COPTR manuscript (Sherwood et al., 2013) prior to coding COPTR intervention materials. Agreement between coding pairs was similar with pair 1 and pair 2 achieving a prevalence- and bias-adjusted kappa (PABAK) of .71 and .72, respectively. External coders had no previous knowledge of COPTR interventions.

Coding Process

External Coders—External coding pairs coded each study's documents prior to internal coders. Each site's materials were coded by a single external coder pair (two individual coders). One pair completed a portion of Minnesota's materials and all Case Western's materials. A second pair completed the second portion of Minnesota's materials and all Vanderbilt's and Stanford's materials. Materials were coded in the following order: Minnesota, Case Western, Vanderbilt, and Stanford. Following instructions for coding

intervention content, coders independently coded by highlighting and adding comments electronically to intervention documents (Michie et al., 2014). Individual coders from a coding pair met three times throughout coding to compare independent coding and create an adjudicated version. External coders met to adjudicate after completing approximately 10%, 50% and 100% of pages in each study's materials. The specific documents included in each portion of the studies' intervention materials are shown in Table 1. External coders developed study-specific coding templates specifying target behaviors (e.g., physical activity, diet, sleep, stress, screen time) and target participants (e.g., parent, child, both).

Internal Coders—A single pair of trained internal coders at each COPTR site coded their own study's materials following the process outlined above for external coders. Each coding pair met after independently coding approximately 10%, 50% and 100% of materials to create an adjudicated version. Internal coders at each study used the template of target behaviors and target participants created by external coders.

Statistical Analysis

Statistical descriptions of data, including agreement charts, describe the number of techniques identified by each set of coders (Bangdiwala, 2016). Inter-coder reliabilities were calculated using PABAK to adjust for chance agreement, prevalence of occurrence, and bias in technique identification rates (Byrt et al., 1993). Inter-coder reliabilities were calculated between internal coders at each site, at specified intervals (i.e. at 10%, 50% and 100% completion milestones during coding) and between internal and external coders. The inter-coder reliabilities between internal and external coders were further examined by target behavior and target participants. To determine whether or not coding was adequately successful, the following standard ranges for interpreting agreement were used: *almost perfect* agreement (.75 to .99); *substantial* agreement (.45 to .74); *moderate* agreement (.20 to .44); *fair* agreement (0.00 to .19); and *poor* agreement (<0.00) (Muñoz & Bangdiwala, 1997). For this study, anything below .45 was considered inadequate.

Results

Aim 1- Agreement Between Internal and External Coders

Coding time ranged from 4 to 22 minutes per page for internal coders. Minnesota took an average of 65 hours per coder to complete 1051 pages of materials (~4 minutes per coder per page) while Case Western took 57 hours per coder for 682 pages (~5 minutes per coder per page). Vanderbilt took 106 hours per coder for 294 pages (~22 minutes per coder per page) and Stanford took 45 hours per coder for 453 pages (~6 minutes per coder per page). On average, internal coders spent 25 hours on each study's intervention materials (~2 minutes per page). Inter-coder reliabilities between internal and external coders indicated *moderate* and *substantial* agreement with PABAK values ranging from .38 to .55 across studies as indicated in Table 2.

Figure 1 shows the patterns of agreement between coders. As shown in Figure 1, bar graph A, the number of techniques congruently identified by both internal and external coders ranged from 31 to 43. Figure 1, bar graph D depicts the number of techniques

congruently not identified by either internal or external coders, ranging from 33 to 36. In Figure 1, bar graph B, the number of techniques identified by external coders only ranged from four to 28 and in Figure 1, bar graph C, the number of techniques identified by internal coders only ranged from one to 23. Minnesota internal coders identified only one technique not also identified by external coders. External coders, however, identified 28 techniques in the Minnesota materials that internal coders did not identify. A similar pattern was observed for Stanford with internal coders identifying only two techniques not identified by external coders and external coders identifying 12 techniques not identified by internal coders. The reverse pattern was observed for Vanderbilt with internal coders identifying 23 techniques not identified by external coders and external coders identifying only four techniques not identified by internal coders. For the Case Western materials, a similar number of techniques were identified by internal coders only (12 techniques) and external coders only (13 techniques).

Sixteen techniques were identified as present by both internal and external coders in all four sites, all of which were included in the online training and are commonly used (e.g., *goal setting*, *problem solving*, and *action planning*). There were 15 techniques not found in any of the materials by internal or external coders, of which only two were covered in the online training. Most were less-common techniques (e.g., *imaginary punishment*, *overcorrection*). There were no commonalities across sites in terms of specific techniques discrepantly coded.

Aim 2 - Reliability Between Internal and External Coders Across Target Behavior and Target Participant

Table 3 presents inter-coder reliabilities between internal and external coders for specific target behaviors (e.g., diet, physical activity, TV or sleep), general health behaviors or other specific non-health-related behaviors. Inter-coder reliability, measured by PABAK ranged from .44 to .63 for diet (*moderate* to *substantial* agreement); .63 to .76 for physical activity (*substantial* to *almost perfect* agreement); .59 to .66 for television (*substantial* agreement; Minnesota and Stanford only); .72 to 1.00 for sleep (*almost perfect* agreement); and .68 to .98 for stress (*substantial* to *almost perfect* agreement; Case Western and Vanderbilt only).

Results by target participant are also shown in Table 3. Reliability between internal and external coder PABAK ranged from .44 to .68 (*moderate* to *substantial* agreement) in study materials that directed the interventionist to work primarily with the parent to influence child behaviors and outcomes. Reliability ranged from .29 to .66 (*moderate* to *substantial* agreement) for study materials that directed the interventionist to work primarily with the child to influence their own behaviors and outcomes. Reliability ranged from .32 to .68 (*moderate* to *substantial* agreement) in study materials that directed the interventionist to work concurrently with the parent and child to influence child behaviors or outcomes. Reliability ranged from .76 to .96 (*almost perfect* agreement) between internal and external coders when the target participant was unclear. Though this was an infrequent occurrence, in the Minnesota materials, external coders only identified four of 93 techniques in study materials that did not have a clear target participant and internal coders did not identify any techniques in study materials without a clear target participant.

Aim 3- Reliability Within Internal and External Coder Pairs at Specified Intervals

Table 4 also presents within-pair reliability at specified intervals. Among internal coders, the mean reliability difference between the first 10% and the last 50% of the intervention materials was .03 (SD = .09). For Minnesota and Case Western internal coder pairs, reliability was in the *substantial agreement* range for the first 10% of the materials and in the *almost perfect* range for 51% –100% intervention category. Reliability estimates for the Vanderbilt internal coder pair was in the *substantial* range for both the first 10% and 51% –100% of the intervention materials. Reliability estimates for the first 10% and the last 51% –100% of intervention materials for the Stanford internal coder pair were in the *almost perfect* agreement range.

Aim 4- Final Adjudication Process Between Internal and External Coders

After calculating reliabilities and evaluating discrepancies between internal and external coders, a final adjudication process was developed. The full set of results for each coding pair and the adjudicated decision by technique can be found in Online Supplementary Table 1. The first step was to identify the discrepant techniques. Excerpts of coded materials corresponding to each discrepant techniques were compiled and compared. Two conferences between COPTR and University College London investigators were conducted by phone to discuss the rationale for coding discrepant techniques. Final decisions were made by consensus regarding the presence or absence of specific techniques for the coded materials. The majority of discrepancies were changed to reflect the original coding of the external coders and were due to different interpretations of the definition of a particular technique. Less often, discrepancies were due to inferences about intervention materials made by the internal coders given their familiarity with the intervention. Additionally, some discrepancies were changed to reflect the original coding of the internal coders, and these discrepancies were usually due to misinterpretation of intervention materials by the external coders.

Discussion

Identifying the behavior change techniques used in effective and ineffective obesity interventions can provide insight into which techniques lead to change and which techniques are unnecessary. To do this work systematically throughout the field, researchers need to be able to reliably identify techniques within their own interventions. Our results largely demonstrate the feasibility of training researchers to code techniques in intervention materials via the online Taxonomy training. The vast majority of calculated inter-coder reliabilities were above the pre-determined adequately successful threshold of .45 (*substantial* agreement or better). However, most of the inter-coder reliability estimates did not achieve a more stringent reliability threshold of 0.75 (*almost perfect* agreement). Further, these measures of reliability in coding of intervention materials between internal and external coders were notably lower than reliability previously reported in external coders coding intervention descriptions (Abraham et al., 2015). Given this, and the final adjudication process, a number of concrete suggestions for future research to improve reliability in coding are provided below.

There were five major takeaways from the adjudication process and resulting recommendations for the field. First, researchers should develop guidance on how to specify techniques in terms of the target participant. There were a number of techniques that were co-delivered to the parent as *Instruction* on how to deliver techniques to their child in order to support the child's behavior change. We recommend that future versions of the Taxonomy include information on how to specify these co-delivered Techniques. This would also aide in distinguishing between techniques such as *Self-Monitoring* and *Monitoring of Behaviors by Others*, which was a common discrepancy. Second, the researchers should develop a set of specific *a priori* decisions about coding prior to initiation to improve reliability and reduce the time needed for coding. These decisions should be published alongside results. Prior to initiating coding, we recommend that all coders identify the intervention's target participant(s), behavior(s), and outcome(s); this may make it easier to distinguish between BCTs such as *Goal Setting (Behavior)* and *Goal Setting (Outcome)*, which was a common discrepancy. Third, researchers should intentionally select which types of materials to code *a priori* so less adjudication may be needed throughout. We recommend intentionally selecting the materials to be coded in the 10%, 50%, 100% adjudication steps to include selections from each type of intervention materials coded (e.g., participant workbooks, training manuals, handouts). Fourth, researchers should develop training materials for coders to support coding intervention materials in addition to manuscript decisions. We recommend training and certification on materials similar to those coded (i.e., intervention materials with similar target participants, behaviors, and outcomes). Lastly, there is a need for intervention developers to better specify their interventions using the Taxonomy. When designing and developing intervention materials, researchers should consider using standardized definitions, such as the Behavior Change Technique Taxonomy, to clearly label proposed active ingredients.

Although many of the techniques identified as present by both internal and external coders were among the 44 considered commonly-used techniques and included in the online training (Michie et al., 2015), there were discrepancies in some techniques that were not included in the training. Further training on the remaining techniques could be warranted if identified as a priority by researchers. Some techniques were discrepantly coded across three or four sites suggesting that further training is needed for the following: *monitoring of behavior by others without feedback*, *information about antecedents*, *review outcome goals*, *feedback on behavior*, *associative learning*, *graded tasks*, *social incentives*, and *framing/reframing*. There did not appear to be a relationship between the amount of time to code or length of materials and reliability. Future research using advanced techniques such as content analysis could be useful in identifying which specific techniques are problematic and why.

During adjudication, the primary noted reasons for discrepancies was due to misinterpretation of the techniques rather than the materials, which supports the inclusion of specific training around reducing bias when coding ones' own intervention materials. Additionally, there was variability in reliability within coding pairs across different intervals of intervention materials (e.g., 1% -10%, 11% -50%, 51% -100%). This is likely due in part to different types of intervention materials included in various intervals. Because not all types of materials are covered equally in the online training with some not covered at

all, additional training modules should be created for various types of intervention materials. Similar findings have been supported elsewhere (Wood et al., 2016). The way individual techniques are described in different types of intervention materials may have restricted coders' abilities to identify techniques. As the importance of using this methodology to improve the description of intervention content becomes more well-known in research and practice communities, more specific language can be used within different types of intervention materials, which may have the additional benefit of improving the specificity with which they are delivered. Decisions made between coder pairs during the adjudication process may lead to more similar results within a coding pair, but less similar results with experienced Taxonomy coders (Abraham et al., 2015). However, improvements in coding within pairs were not seen consistently across sites. It is possible that different patterns may have been related to the variability in coders' familiarity with the techniques, their role in the interventions, or their background and related education.

Given the potential benefits of using a common language to identify techniques designed within interventions, steps were taken to understand the discrepancies between internal and external coders in terms of target behaviors and target participants. Similar to target behaviors, coders are also required to identify the target "outcome" to correctly code certain techniques. This source of discrepancy became clear during the final adjudication process. To correctly adjudicate, the decision was made to identify only one target outcome for each of these trials, either a reduction in weight status or body mass index trajectory. But certain target behaviors (e.g., child diet) could have been identified as intermediate outcomes of other target behaviors (e.g., diet-related parenting practices). Taxonomies and ontologies of target behaviors are currently being developed to help researchers use consistent terminology and may improve the reliability of coding efforts such as these (Larsen et al., 2017). Although the range of inter-coder reliability estimates was similar regardless of the target participant, the final adjudication process identified this a key area for future research. In behavior change interventions targeting child outcomes (e.g., childhood obesity), parents were often targeted as intermediaries. Similarly, although there were no obviously different patterns of techniques in the studies for older children compared to younger children, some techniques may not be relevant when only targeting younger children. For example, goal setting around weight loss outcomes is not an appropriate technique to use with preschoolers. Making *a priori* decisions about which techniques could apply by target participant or age could help improve reliability by reducing the total number of techniques from which to choose. Further work needs to be done to determine which techniques can be delivered to which participants and how co-delivered techniques should be classified.

The strengths of this study most notably include: the novelty of specifying intervention content across different interventions with the same outcome, using robust measures of reliability (i.e., PABAK and agreement charts), and the development of a formal adjudication process to identified areas for improvement. Limitations of this study included the inability to capture "frequency" or "dose" of techniques in a meaningful way as well as the inability to adjust analyses for potential clustering or non-independence of techniques. In addition, this study focused only on techniques designed rather than those delivered, received or enacted by participants (i.e., treatment fidelity). In order to report fidelity to

an intervention, the intended intervention must first be specified. This project serves as the groundwork for future research testing whether or not interventions are delivered as specified in the design. Although developing methods for identifying intervention dose, clustering and treatment fidelity are warranted (Lorencatto et al., 2014), standardized and feasible approaches to conducting such analyses do not yet exist. Because this project aimed to serve as a feasible example of how researchers can code techniques in their own studies, these advanced approaches were not taken here.

Though this project focused on coding of techniques in childhood obesity interventions, these results and approaches from this study have applicability more broadly across health promotion interventions. This Taxonomy and online coder training can be used by health promotion practitioners to better understand the discrete components of behavior change and how these align with behavior change theory and theoretical frameworks. This would allow practitioners to design or redesign interventions in more targeted, efficient way. The current project is an exciting step in the application of the Taxonomy framework to behavior change interventions for pediatric obesity in that it provides a necessary next step to help researchers identify which components of an intervention were or were not effective and why.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding Information

This research was his research is being supported by the National Institutes of Health grant numbers: U01 HL103561, U01 HL103620, U01 HL103622, U01 HL103629, and U01 HD068890. Trials were registered on [ClinicalTrials.gov](https://clinicaltrials.gov) (NCT01606891, NCT01514279, NCT01316653, and NCT01642836). The content expressed in this paper is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, And Blood Institute, the Eunice Kennedy Shriver National Institute of Child Health and Human Development, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Biographies

Meghan M. JaKa, PhD is an Evaluation Associate at the HealthPartners Institute, Center for Evaluation and Survey Research, Minneapolis, MN, US with no potential conflicts of interest to declare.

Caroline E. Wood, PhD is an Honorary Senior Researcher at the Centre for Digital Public Health in Emergencies, University College London, UK with no potential conflicts of interest to declare.

Sara Veblen-Mortenson, MPH/MSW is a Program Manager in the Dean's Office, School of Public Health at the University of Minnesota, MN, US with no potential conflicts of interest to declare.

Shirley M. Moore, PhD, RN, is the Edward J. and Louise Mellen Professor of Nursing Emerita; Distinguished University Professor Emerita at the Frances Payne Bolton School of Nursing, Case Western Reserve University, Cleveland, OH, US with no potential conflicts of interest to declare.

Donna M Matheson, PhD is Senior Research Scientist in the Department of General Pediatrics, Stanford Medical School, Stanford University, Stanford, CA, US with no potential conflicts of interest to declare.

June Stevens, PhD is a Distinguished Professor in the Departments of Nutrition and Epidemiology at the University of North Carolina, Chapel Hill, NC, US and receives support from a research grant to the University from Weight Watchers.

Lou Atkins, PhD is Senior Teaching Fellow at Centre for Behaviour Change, University of London, UK with no potential conflict of interest to declare.

Susan Michie BA, MPhil, DPhil is Professor of Health Psychology and Director of the Centre for Behaviour Change at University College London, UK, with no potential conflicts of interest to declare.

Clara Adegbite-Adeniyi, BA is a Research Project Manager at Case Western Reserve University, Frances Payne Bolton School of Nursing, Cleveland, OH, US with no potential conflicts of interest to declare.

Oluwatomisin Olayinka, PhD,RN, MPH is an Assistant Professor at MGH Institute of Health Professions School of Nursing, Boston, MA, US with no potential conflict of interest to declare.

Eli K. Po'e, BS is a Research Assistant with Vanderbilt University Medical Center, Nashville, TN, US with no potential conflicts to disclose.

Alethea M Kelly, BS is a Budget Accounting Analyst Pre-Award, Vanderbilt University Medical Center, Nashville, TN, US with no potential conflicts of interest to declare.

Holly Nicastro, PhD, MPH, is a Program Director in the Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, Bethesda, MD, US with no potential conflicts of interest to declare.

Shrikant I. Bangdiwala, PhD is Professor in the Department of Health Research Methods, Evidence and Impact, and Director of Statistics at the Population Health Research Institute, McMaster University, Hamilton, ON, CA, with no potential conflicts of interest to declare.

Shari Barkin, MD, MSHS is a Professor of Pediatrics at Vanderbilt University Medical Center, Nashville, TN, US with no potential conflicts of interest to declare.

Charlotte A. Pratt, PhD, RD is a Health Scientist Administrator and Deputy Branch Chief at the National Heart, Lung, and Blood Institutes, National Institutes of Health, Bethesda, MD, US with no conflicts of interests to declare.

Thomas N. Robinson, MD, MPH is the Irving Schulman, MD Endowed Professor in Child Health, Professor of Pediatrics and of Medicine, and Director of the Stanford Solutions

Science Lab and the Center for Healthy Weight at Stanford University and Lucile Packard Children's Hospital Stanford, Stanford, CA, US with no potential conflicts to report.

Nancy E. Sherwood, PhD is an Associate in the Division of Epidemiology and Community Health, School of Public Health, Minneapolis, MN, US with no potential conflicts of interest to declare.

References

- Abraham C, Wood CE, Johnston M, Francis J, Hardeman W, Richardson M, & Michie S (2015). Reliability of identification of behavior change techniques in intervention descriptions. *Annals of Behavioral Medicine*, 49(6), 885–900. 10.1007/s12160-015-9727-y [PubMed: 26290002]
- Bandura A (1997). *Self-efficacy: the exercise of control*. W. H. Freeman & Company.
- Bangdiwala SI (2016). Graphical aids for visualizing and interpreting patterns in departures from agreement in ordinal categorical observer agreement data. *Journal of Biopharmaceutical Statistics*, 27(5), 773–783. 10.1080/10543406.2016.1273941
- Bangdiwala SI, Bhargava A, O'Connor DP, Robinson TN, Michie S, Murray DM, Stevens J, Belle SH, Templin TN, & Pratt CA (2016). Statistical methodologies to pool across multiple intervention studies. *Translational Behavioral Medicine*, 6(2), 228–235. 10.1007/s13142-016-0386-8 [PubMed: 27356993]
- Barkin SL, Heerman WJ, Sommer EC, Martin NC, Buchowski MS, Schlundt D, Po'e EK, Burgess LE, Escarfuller J, & Pratt C (2018). Effect of a behavioral intervention for underserved preschool-age children on change in body mass index: A randomized clinical trial. *JAMA*, 320(5), 450–460. 10.1001/jama.2018.9128 [PubMed: 30088008]
- Black N, Johnston M, Michie S, Hartmann-Boyce J, West R, Viechtbauer W, Eisma MC, Scott C, & de Bruin M (2020). Behaviour change techniques associated with smoking cessation in intervention and comparator groups of randomized controlled trials: A systematic review and meta-regression. *Addiction*. 10.1111/add.15056
- Byrt T, Bishop J, & Carlin JB (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology*, 46(5), 423–429. 10.1016/0895-4356(93)90018-V [PubMed: 8501467]
- Dweck CS (2017). The journey to children's mindsets-and beyond [Peer Reviewed]. *Child Development Perspectives*, 11(2), 139–144. 10.1111/cdep.12225
- French SA, Sherwood NE, Veblen-Mortenson S, Crain AL, JaKa MM, Mitchell NR, Hotop AM, Berge JM, Kunin Batson AS, & Truesdale K (2018). Multicomponent obesity prevention intervention in low-income preschoolers: Primary and subgroup analyses of the NET-Works randomized clinical trial, 2012–2017. *American Journal of Public Health*, 108(12), 1695–1706. 10.2105/AJPH.2018.304696 [PubMed: 30403521]
- Larsen KR, Michie S, Hekler EB, Gibson B, Spruijt-Metz D, Ahern D, Cole-Lewis H, Ellis RJ, Hesse B, Moser RP, & Yi J (2017). Behavior change interventions: The potential of ontologies for advancing science and practice. *Journal of Behavioral Medicine*, 40(1), 6–22. 10.1007/s10865-016-9768-0 [PubMed: 27481101]
- Lorenatto F, West R, Bruguera C, & Michie S (2014). A method for assessing fidelity of delivery of telephone behavioral support for smoking cessation. *Journal of Consulting and Clinical Psychology*, 82(3), 482–491. [PubMed: 24294836]
- Martin J, Chater A, & Lorenatto F (2013). Effective behaviour change techniques in the prevention and management of childhood obesity. *International Journal of Obesity*, 37(10), 1287–1294. 10.1038/ijo.2013.107 [PubMed: 23756676]
- Michie S, Atkins L, & West R (2014). *The Behaviour Change Wheel: A Guide to Designing Interventions* (1st ed.). Silverback Publishing.
- Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, Eccles MP, Cane J, & Wood CE (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine*, 46(1), 81–95. 10.1007/s12160-013-9486-6 [PubMed: 23512568]

- Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, & Hardeman W (2015). 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). *Health Technology Assessment*, 19(99), 1–188. 10.3310/hta19990
- Moore SM, Borawski EA, Cuttler L, Ievers-Landis CE, & Love TE (2013). IMPACT: A multi-level family and school intervention targeting obesity in urban youth. *Contemporary Clinical Trials*, 36(2), 574–586. 10.1016/j.cct.2013.08.009 [PubMed: 24008055]
- Moore SM, Borawski EA, Love TE, Jones S, Casey T, McAleer S, Thomas C, Adegbite-Adeniyi C, Uli NK, & Hardin HK (2019). Two family interventions to reduce BMI in low-income urban youth: A randomized trial. *Pediatrics*, 143(6), 1–11. 10.1542/peds.2018-2185
- Muñoz SR, & Bangdiwala SI (1997). Interpretation of Kappa and B statistics measures of agreement. *Journal of Applied Statistics*, 24(1), 105–112. 10.1080/02664769723918
- Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, Mullany EC, Biryukov S, Abbafati C, Abera SF, Abraham JP, Abu-Rmeileh NM, Achoki T, AlBuhairan FS, Alemu ZA, Alfonso R, Ali MK, Ali R, Guzman NA, ... Gakidou, E. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet*, 384(9945), 766–781. 10.1016/S0140-6736(14)60460-8 [PubMed: 24880830]
- Po'e EK, Heerman WJ, Mistry RS, & Barkin SL (2013). Growing Right Onto Wellness (GROW): A family-centered, community-based obesity prevention randomized controlled trial for preschool child-parent pairs. *Contemporary Clinical Trials*, 36(2), 436–449. 10.1016/j.cct.2013.08.013 [PubMed: 24012890]
- Pratt CA, Boyington J, Esposito L, Pemberton VL, Bonds D, Kelley M, Yang S, Murray D, & Stevens J (2013). Childhood Obesity Prevention and Treatment Research (COPTR): Interventions addressing multiple influences in childhood and adolescent obesity. *Contemporary Clinical Trials*, 36(2), 406–413. 10.1016/j.cct.2013.08.010 [PubMed: 23999502]
- Psaltopoulou T, Tzanninis S, Ntanasis-Stathopoulos I, Panotopoulos G, Kostopoulou M, Tzanninis I-G, Tsagianni A, & Sergentanis TN (2019). Prevention and treatment of childhood and adolescent obesity: A systematic review of meta-analyses. *World Journal of Pediatrics*, 1–32. 10.1007/s12519-019-00266-y
- Resnicow K, Baranowski T, Ahluwalia JS, & Braithwaite RL (1999). Cultural sensitivity in public health: Defined and demystified. *Ethnicity & Disease*, 9(1), 10–21. <https://www.ncbi.nlm.nih.gov/pubmed/10355471> [PubMed: 10355471]
- Robinson TN (2010). Save the world, prevent obesity: Piggybacking on existing social and ideological movements. *Obesity*, 18 (Suppl 1), S17–22. 10.1038/oby.2009.427 [PubMed: 20107456]
- Robinson TN, Kraemer HC, Matheson DM, Obarzanek E, Wilson DM, Haskell WL, Pruitt LA, Thompson NS, Haydel KF, Fujimoto M, Varady A, McCarthy S, Watanabe C, & Killen JD (2007). Stanford GEMS phase 2 obesity prevention trial for low-income African-American girls: Design and sample baseline characteristics. *Contemporary Clinical Trials*, 29(1), 56–69. 10.1016/j.cct.2007.04.007 [PubMed: 17600772]
- Robinson TN, Matheson D, Desai M, Wilson DM, Weintraub DL, Haskell WL, McClain A, McClure S, J, A. B., Sanders, L. M., Haydel, K. F., & Killen, J. D. (2013). Family, community and clinic collaboration to treat overweight and obese children: Stanford GOALS-A randomized controlled trial of a three-year, multi-component, multi-level, multi-setting intervention. *Contemporary Clinical Trials*, 36(2), 421–435. 10.1016/j.cct.2013.09.001 [PubMed: 24028942]
- Sherwood NE, French SA, Veblen-Mortenson S, Crain AL, Berge JM, Kunin-Batson A, Mitchell N, & Senso M (2013). NET-Works: Linking families, communities and primary care to prevent obesity in preschool-age children. *Contemporary Clinical Trials*, 36(2), 544–554. 10.1016/j.cct.2013.09.015 [PubMed: 24120933]
- Tate DF, Lytle LA, Sherwood NE, Haire-Joshu D, Matheson D, Moore SM, Loria CM, Pratt C, Ward DS, Belle SH, & Michie S (2016). Deconstructing interventions: Approaches to studying behavior change techniques across obesity interventions. *Translational Behavioral Medicine*, 6(2), 236–243. 10.1007/s13142-015-0369-1 [PubMed: 27356994]

- Wood CE, Hardeman W, Johnston M, Francis J, Abraham C, & Michie S (2016). Reporting behaviour change interventions: Do the behaviour change technique taxonomy v1, and training in its use, improve the quality of intervention descriptions? *Implementation Science*, 11(1), 84. 10.1186/s13012-016-0448-9 [PubMed: 27268131]
- Wood CE, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, & Michie S (2015). Applying the Behaviour Change Technique (BCT) Taxonomy v1: A study of coder training. *Translational Behavioral Medicine*, 5(2), 134–148. 10.1007/s13142-014-0290-z [PubMed: 26029276]

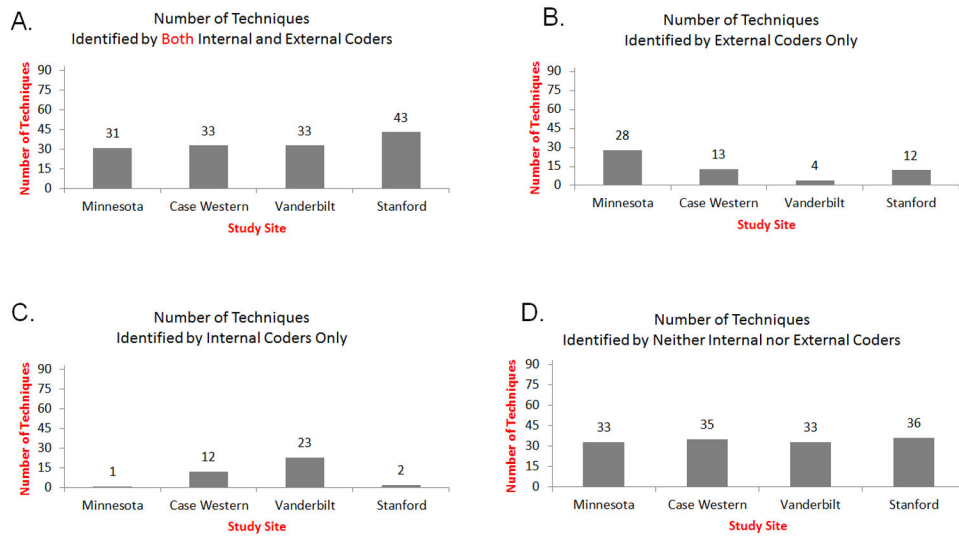


Figure 1. Number of Techniques Identified by Coding Pairs Across Sites
Note: Bar graphs A-D depict the patterns of agreement between coders.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Specific Documents Included in the Intervention Materials Coded by Internal and External Coders

Study Site	Documents		
	1% - 10% Materials	11% - 50% Materials	51% - 100% Materials
Minnesota ¹	Parent Educator manual (Sessions 1-3)	Parent Educator manual (Sessions 4-14)	Family Connector manual, activity cards, tip sheets, sample packet, primary care provider training presentation
Case Western ¹	Intervention manual, handouts (Sessions 1-3)	Intervention manual, handouts (Sessions 4-11)	Intervention manual, handouts (Sessions 12-25)
Vanderbilt ²	Intervention booklets (Sessions 1-4)	Intervention booklets (Sessions 4-12), intervention facilitators guide	Intensive phone call scripts, maintenance phase phone call scripts, adaptive intervention phone call scripts, body mass index feedback reports
Stanford ²	Home visiting manual, handouts (Diet Lessons 1-4)	Home visiting manual, handouts (Diet Lessons 5-7, Environmental Changes Lessons 1-4 and Physical Activity Lessons 1-5)	Home visiting manual, handouts (Physical Activity Lessons 6-7, Screen Time Lessons), handouts, and afterschool sports curriculum

¹Coded by study specific internal coding pair and external coding pair 1

²Coded by study specific internal coding pair and external coding pair 2

Table 2

Inter-coder Reliability Between Internal Coders and External Coders Within Each Study Site's Intervention Materials

Variable	Study Site			
	Minnesota	Case Western	Vanderbilt	Stanford
PABAK	.38	.46	.41	.55
Number of techniques identified by external coders	59	46	37	48
Number of techniques identified by internal coders	32	45	56	45

Note: PABAK = Prevalence- and Bias-adjusted Kappa

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Inter-coder Reliability Between External Coders and Internal Coders by Target Behavior and Target Participant

Study Site	Inter-coder Reliability (PABAK)						
	Target Behavior						
	Diet	Physical Activity	TV	Sleep	Stress	General ¹	Other ²
Minnesota	.63	.68	.66	.94	n/a	0.00	.40
Case Western	.58	.76	n/a	.70	.68	.42	.85
Vanderbilt	.44	.74	n/a	.72	.98	.61	.46
Stanford	.51	.63	.59	1.00	n/a	.81	.68
	Target Participant						
	Parent	Child	Both	Unclear ³			
Minnesota	.68	.46	.68	.91			
Case Western	.53	.29	.51	.76			
Vanderbilt	.44	.66	.32	.96			
Stanford	.63	.33	.43	.94			

Note: PABAK = Prevalence and Bias-adjusted Kappa

¹The behavior was unspecified, typically with regards to healthy behaviors in general

²The behavior was specified but not an *a priori* identified health behavior (e.g., mindfulness)

³It was unclear whether the coded technique was specific to the parent, child or both

Table 4

Inter-coder Within-pair Reliability (PABAK) for Internal and External Coders at Specified Intervals

Coder Pairs	Study Site			
	Minnesota	Case Western	Vanderbilt	Stanford
Internal Coders				
1% -10%	.74	.68	.63	.83
11% -50%	.74	.54	.44	.81
51% -100%	.83	.80	.59	.76
External Coders (Pair 1)				
1% -10%	.69	.67	-	-
11% -50%	.76	.85	-	-
51% -100%	.81	.83	-	-
External Coders (Pair 2)				
1%-10%	-	-	.85	.85
11%-50%	-	-	.81	.85
51%-100%	-	-	.85	.94

Note. PABAK = Prevalence and Bias-adjusted Kappa

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript