9-2019

Development of a Novel Tool for Assessing Coverage of Implementation Factors in Health Promotion Program Resources

Hannah Calvert
Boise State University, hannahcalvert898@boisestate.edu

Lindsey Turner
Boise State University, lindseyturner1@boisestate.edu

Publication Information

Calvert, Hannah and Turner, Lindsey. (2019). "Development of a Novel Tool for Assessing Coverage of Implementation Factors in Health Promotion Program Resources". *Preventative Medicine Reports, 15*, 100909-1 100909-5. https://dx.doi.org/10.1016/j.pmedr.2019.100909

Preventive Medicine Reports 15 (2019) 100909

Contents lists available at ScienceDirect

Preventive Medicine Reports

journal homepage: www.elsevier.com/locate/pmedr

Development of a novel tool for assessing coverage of implementation factors in health promotion program resources

Carolina M. Bejarano[a,b], Kelli Snow[b], Hannah Lane[c], Hannah Calvert[d], Kate Hoppe[b], Nicole Alfonsin[e], Lindsey Turner[d], Jordan A. Carlson[b,*]

[a] Clinical Child Psychology Program, University of Kansas, 1000 Sunnyside Avenue, Lawrence, KS, United States of America
[b] Center for Children's Healthy Lifestyles and Nutrition, Children's Mercy Kansas City, 610 E. 22nd St, Kansas City, MO, United States of America
[c] Department of Pediatrics, Division of Growth and Nutrition, University of Maryland School of Medicine, 737 W Lombard St, Baltimore, MD, United States of America
[d] College of Education, Boise State University, 1910 University Drive, Boise, ID, United States of America
[e] George Washington University Milken Institute, School of Public Health, 950 New Hampshire Avenue NW, Washington, DC, United States of America

1. Introduction

Programs for promoting health behaviors have been developed for adoption in various settings such as schools (Hoelscher et al., 2010; Sallis et al., 1997), afterschool programs (Beets et al., 2009), child care centers (Ward et al., 2010), community health and mental health clinics (Aarons et al., 2011), and other settings (Soler et al., 2010). While many of these programs have established efficacy, implementation-related challenges often limit their impact in real-world applications (Glasgow et al., 2003; Glasgow and Emmons, 2007). As a result, an emerging body of literature has pointed to the need to better consider implementation factors related to real-world applications when developing setting-based health promotion programs (Tomoaia-Cotisel et al., 2013; Damschroder et al., 2009). Also of critical importance is to prioritize the dissemination of programs that address implementation factors, and to supplement programs that do not address such factors with appropriate and effective supports, to maximize implementation rates and quality.

Implementation factors are the elements that surround a particular intervention, which impact the extent to which it is put into practice as intended, and the extent to which it is successful in promoting the intended outcomes (Damschroder et al., 2009; Rabin et al., 2008). These factors span multiple levels of influence, in accordance with ecological models, including individual-, interpersonal-, and environmental-level influences (Sallis et al., 2015). A growing body of evidence shows the importance of such factors to ensure "real-world" program effectiveness–even the most efficacious programs will not work without quality implementation (Tomoaia-Cotisel et al., 2013; Durlak and DuPre, 2008). Several implementation science frameworks have been developed to organize the conceptualization of implementation factors (Tabak et al., 2012a; Nilsen, 2015). Broadly, these frameworks include characteristics of the intervention, social and physical environment, availability of resources, and characteristics of the individuals involved in implementation. While implementation science frameworks have traditionally been used to identify barriers and facilitators to implementation and/or guide implementation efforts (Nilsen, 2015; Tabak et al., 2012b), they should also be used to inform the selection, prioritization, refinement, and use of available programs or program components to maximize implementation rates, quality, and effectiveness.

Tools and methods have been developed and applied to rate policies, practices, and programs on characteristics such as their quality and effectiveness (Schwartz et al., 2009; Collins et al., 2002; Carlson et al., 2013). However, tools for assessing implementation factors within available health promotion programs have been given little attention. Therefore, the purpose of the present study was to develop and test a tool and methodology for assessing coverage of a wide range of theoretically-based implementation factors within available health promotion programs. The study was informed by the Consolidated Framework for Implementation Research (CFIR) (Damschroder et al., 2009), which is a consolidation of multiple frameworks and has been used across a variety of studies (Kirk et al., 2016). The tool involves a coding system that allows researchers and other stakeholders to review and select programs to support their mission of health promotion, with key consideration of each program's potential for optimal implementation, as a result of the types of implementation supports that accompany each program. This paper describes the process of developing and testing a tool specific to classroom-based physical activity (CBPA) programs, with methodology that can be applied to a wide range of health promotion programs, particularly those that are setting- or organization-based.

* Corresponding author at: Center for Children's Healthy Lifestyles and Nutrition, Children's Mercy Kansas City, 610 E. 22nd St, Kansas City, MO 64113, United States of America.
E-mail address: jacarlson@cmh.edu (J.A. Carlson).

https://doi.org/10.1016/j.pmedr.2019.100909
Received 28 September 2018; Received in revised form 22 May 2019; Accepted 26 May 2019
Available online 30 May 2019

2. Methods

2.1. Study overview

This study involved the development and testing of a coding tool and methodology for assessing whether and to what extent available CBPA programs address factors that relate to improved implementation in "real-world" settings. To accomplish this, a team of experts with experience in CBPA research was organized. The team selected an implementation science framework to guide the study, developed a set of codes specific to the implementation factors, developed code definitions and a coding guide, systematically identified and coded available CBPA programs, tested the quality of the codes by having two coders independently code each program, and subsequently calculated inter-rater reliability.

2.2. Selection of implementation framework

CFIR was selected to guide the development of the implementation factor assessment tool because of its comprehensiveness (Damschroder et al., 2009) relevance to implementation effectiveness in previous studies (Kirk et al., 2016), and coverage of implementation factors acting at multiple ecological levels of influence. CFIR includes 36 constructs organized under five domains: Intervention Characteristics (key attributes of interventions influence the success of implementation), Outer Setting (factors outside of an organization that affect implementation), Inner Setting (factors within an organization that affect implementation), Characteristics of Individuals (perceptions, attitudes, and motivation of individuals within the organization), and Process (strategies used to support implementation) (Damschroder et al., 2009). Other frameworks (Tabak et al., 2012a) were considered and would have been appropriate (Atkins et al., 2017; Kitson et al., 2008; Aarons et al., 2011). Ultimately, CFIR was chosen over these frameworks because of the team's familiarity with the framework and the extensive resources and descriptions available on the CFIR website (Consolidated Framework for Implementation Research, 2014).

2.3. Development of codes and code definitions

The CFIR construct list was used to facilitate the development of the list of CBPA-specific implementation factor "codes" (e.g., Goal Setting) and code definitions. The focus was on implementation rather than adoption because implementation is an ongoing effort that is more closely linked with effectiveness than adoption (Rabin et al., 2008). A brief review of the literature was conducted to understand published barriers and facilitators to CBPA implementation. An iterative approach involving all authors was used to develop the codes and corresponding definitions and match them with the CFIR constructs. The primary investigator and a research assistant first drafted the initial code names and definitions, mapped to CFIR. Each CBPA-specific code, definition, and related CFIR construct were then discussed during a series of meetings with the other study team members. During and after each call, the code names, definitions, and mapping to CFIR were refined, and codes were added and removed. After the fourth meeting, consensus was reached and the code list was finalized.

The code development process aimed to produce a tool with code names and definitions that aligned to CFIR, but were more narrowly focused to the school setting and to CBPA programs. Specific differentiating criteria were used to facilitate coding decisions and to maximize relevance to CBPA and minimize subjectivity when applying the codes.

It was acknowledged that various levels of influence affect implementation, and because many CBPA program materials were believed to focus more on the teacher-student relationship (e.g., influencing student behavior change) and less on the school-teacher relationship (e.g., influencing teacher behavior change), each of the agreed upon CBPA-specific codes was categorized in this way. This grouping excluded the CFIR Intervention Characteristics codes because they were not specific to either relationship.

2.4. Identifying programs

A systematic search of both scientific and grey literature was conducted in 2017 to identify programs directed at delivering physical activity in the classroom. It was important to search outside of scientific search engines because health promotion programs are often developed and/or disseminated by community organizations, health organizations, and businesses. Such programs may not have been research-tested but were often research-based, and in some cases (e.g., with regards to CBPA) may be more likely to be adopted than researcher-developed programs. Search terms included classroom physical activity + programs + interventions; active classroom; activity breaks; active lessons; classroom breaks + exercise; classroom physical activity. Search engines included PubMed and Google Scholar for scientific publications, and Google for grey literature. Using Google, the first 50 data hits were examined for each search term. In addition, reference lists of identified scientific publications and programs materials were reviewed for potentially relevant programs or studies. Most programs had their own website but some were housed on another website such as one hosted by a department of education, health department, or health organization.

Potential programs were screened for inclusion as they were identified. Inclusion criteria were: published in English language; targeted increasing physical activity in the classroom (but could also address other sources of in-school physical activity [e.g., Physical Education]); and targeted any/all of grades K-6th. Exclusion criteria for programs were: exclusively offered content designed only for physical education class, recess or afterschool programs, not classroom physical activity; created as part of a research effort but were not publicly available even upon request; previously available, but no longer had active websites or links; and were based outside of the United States.

2.5. Creating directory of program materials

One staff member compiled all materials for each program. First, each program's webpage was visited and thoroughly searched. Downloadable materials (e.g., PDFs, PowerPoints, manuals/guides, handouts, posters) were obtained and compiled in an electronic library. In circumstances when the website mentioned that trainings, materials, or resources were available but were not posted on the website, the staff member contacted the program directly to request these materials. In some cases, website user accounts were created or materials were purchased. A summary document was created for each program, which included an overview of the program and details about its materials, and notes regarding materials that were available through direct contact with the program or on the program's website.

2.6. Coding procedures

A coding guide was created to facilitate the coding of each CBPA program. The coding guide included the CBPA-specific code names and definitions, as well as the related CFIR construct name and short description. Coders were instructed to review all program materials compiled in the library and thoroughly search and review the program's website to obtain information potentially missed by the staff member who initially compiled the resources. Coders were also instructed to exclude program content related to something other than CBPA. This was relevant because several programs addressed physical activity and/or nutrition during other aspects of the school day (e.g., during Physical Education, after school) in addition to CBPA. The team tried to consider how a school or teacher would access/use the resources if they wanted to only implement the CBPA portion of the program.

The CBPA-specific codes within the Intervention Characteristics domain were objectively based, relatively easily identifiable, and the coding required little-to-no judgment or interpretation from the coder. A score option of 0 = no vs. 1 = yes was used for these codes and they were coded by a single coder because double coding was not expected to improve the accuracy of the scores. All other CBPA-specific codes were coded as: 0 = not covered in the program, 1 = covered in the program but at a minimal level, and 2 = covered more than minimally in the program. For these codes, two coders independently coded each program and discrepancies were reconciled through discussion among the two coders. In general, 2's were only used to identify and reflect programs that thoroughly addressed a given concept.

For training, all coders independently coded the same three programs, one at a time, with discussion after each focused on improving inter-rater reliability for the subsequent program coding. A pool of five coders was used. One coder coded all Intervention Characteristics, with all other domains double coded by 2 of the 5 coders (with an approximately equal number of programs being coded by each). Coders were a mix of Masters-level staff and students, PhD students, and postdoctoral fellows with training in research and health promotion.

2.7. Statistical analyses

Inter-rater agreement was assessed using intraclass correlation coefficients (ICCs) and percent agreement, for those coded twice. ICCs were calculated for the 0–2 scores using one-way random effects models for the average of the measures. Percent agreement was calculated for the 0–2 scores. In addition, because coding of "2" was very rare (only 6.7% of codes received any score of 2, across all programs), additional analyses were conducted with codes of 1 and 2 collapsed to 1, versus 0. This dichotomization allowed for examination of whether inter-rater reliability was improved when considering "any versus none" regarding the inclusion of each implementation factor, by grouping the minimally (Durlak and DuPre, 2008) and more than minimally (Carlson et al., 2013) covered scores. ICC magnitude was classified using criteria of poor ($\leq$0.40), fair (0.41–0.60), good (0.61–0.80), and excellent (0.81–1.0) (Cicchetti, 1994) and percent agreement was evaluated with the criteria of $\geq$75% as good to excellent, 60–74% as moderate, and $<$ 60% as poor (Landis and Koch, 1977). Both metrics were considered while interpreting inter-rater reliability, as ICC values can falsely suggest poor agreement when there is little variability in responses across programs. To summarize the responses within the CFIR domains of Inner Setting, Characteristics of Individuals, and Process, and the groupings by the school-teacher, and teacher-student relationship, "index scores," or sum scores within the domains, were calculated. Inter-rater reliability of the index scores was tested using one-way random effects ICCs. Analyses were conducted in SPSS version 23.

3. Results

3.1. Number of programs identified and final list of codes

A total of 37 programs were identified through the search process. The final coding guide included 51 CBPA-specific codes. The code names, code definitions, and mappings to CFIR are presented in the Appendix. Some CFIR constructs were mapped to multiple CBPA-specific codes, whereas others were not mapped to any CBPA-specific constructs, based on the study team's perception of the importance and relevance of the construct to CBPA. This process resulted in 23 CBPA-specific codes for the Intervention Characteristics domain of CFIR, which were coded 0–1 by a single coder, and the other three domains that were coded 0–2 by two coders, including 14 codes for the Inner Setting, 6 for the Characteristics of Individuals, and 8 for Process. The Outer Setting domain was not covered because the codes developed for this domain overlapped with other domains and constructs (e.g., the External Change Agents Construct was included instead in the Process domain), so these codes were mapped to these other constructs. The school-teacher relationship was represented in 21 codes, and the teacher-student relationship in the remaining 7 codes. Across the 28 codes that were scored on the 0–2 scale, 6.7% were scored as 2. Across the CFIR domains, several CFIR constructs were not covered (CFIR Research Team, 2015), either because they were thought to overlap with other constructs, they were considered not be to relevant to CPBA programs, or they were not expected to differ across programs.

Table 1
Rater agreement for implementation factors coded by two coders (N = 37 programs).

Construct	ICC for 3-level score	% agreement for 3-level score	% agreement for dichotomized score
Inner setting			
Communication[a]	0.630	81.1%	83.8%
Policy incorporation[a]	−0.023	64.9%	63.9%
Marketing materials teachers[a]	0.135	56.8%	62.2%
Marketing materials students/parents[b]	0.650	73.0%	86.5%
Gauging/affecting climate[a]	0.854	89.2%	94.6%
Leadership initial buy in[a]	0.488	70.3%	75.7%
Student management[b]	0.729	78.4%	86.5%
Compatibility adaptations[b]	0.539	56.7%	70.3%
Incentives[a]	−0.130	67.6%	67.6%
Goal setting[a]	0.242	86.5%	86.5%
Monitoring[a]	0.679	75.0%	75.7%
School readiness[a]	−0.091	89.2%	89.2%
Leadership engagement post adoption[a]	0.366	86.5%	86.5%
Classroom structure[b]	0.396	62.2%	70.3%
Characteristics of individuals			
Health benefits[a]	0.637	51.3%	67.6%
Non-health benefits[a]	0.695	64.8%	89.2%
Teacher motivation/ attitudes around program[a]	0.036	75.7%	75.7%
Self-efficacy[a]	0.100	75.7%	75.7%
Teacher stage of change[a]	−0.029	94.6%	94.6%
Teacher attitude/value toward PA[a]	0.088	78.4%	78.4%
Process			
Scheduling materials[b]	0.822	73.0%	81.1%
Dose/dose quantity[b]	0.775	72.9%	86.5%
Teacher participation[b]	0.796	75.7%	81.1%
Implementation leaders[a]	0.615	73.0%	73.0%
External involvement[a]	0.467	73.0%	78.4%
External Information sharing[a]	0.409	56.8%	64.9%
Accountability[a]	0.795	97.3%	97.3%
Outcomes[a]	0.704	73.0%	81.1%

[a] Construct affects school-teacher relationship for implementation (21 constructs).
[b] Construct affects teacher-student relationship for implementation (7 constructs).

3.2. Inter-rater agreement for CBPA codes

Inter-rater agreement results (prior to reconciliation) are detailed in Table 1. Fourteen of the 28 codes had excellent inter-rater reliability based on an ICC $>$ 0.80 and/or a percent agreement $\geq$75%. Six codes had good reliability based on an ICC between 0.61 and 0.80. Five codes had moderate inter-rater reliability based on a percent agreement between 60 and 74.9% and an ICC between 0.41 and 0.60. The remaining three codes had poor agreement. Percent agreement for these codes was 57% and ICCs were fair (0.41–0.60) or poor ($\leq$0.40). After all codes were dichotomized as 0 vs. 1, 20 of the 28 codes had good to excellent percent agreements (above 75%) and the remaining 8 had moderate

Table 2
Inter-rater reliability for index scores.

Index	ICC[a]	ICC[b]
Inner setting (14 codes)	0.606	0.592
Characteristics of individuals (6 codes)	0.694	0.614
Process (8 codes)	0.804	0.765
School-teacher relationship (21 codes)	0.681	0.638
Teacher-student relationship (7 codes)	0.847	0.877

[a] For index scores that comprise a sum of the 3-level scores for each code.
[b] For index scores that comprise a sum of the dichotomized scores for each code.

percent agreements (60–74.9%).

3.3. Inter-rater reliability for index scores

Table 2 presents results of the inter-rater reliability analyses pertaining to the multi-code indices. ICCs for the indices calculated from the codes scored 0–2 were all good or excellent, ranging from 0.606 to 0.847. ICCs for the four of the five indices calculated from the codes scored 0–1 were good or excellent, ranging from 0.614 to 0.877. The ICC for the remaining index (Inner Setting) was fair (ICC = 0.592).

4. Discussion

This study presents the process of developing and testing a theoretically-based tool and methodology to assess coverage of implementation factors in setting- and organization-based health promotion programs that have already been developed and disseminated for use. Existing programs often include intervention strategies and content focused on delivery of the intervention to the end target (e.g., delivering CBPA to students). However, little is known about how these programs address implementation factors to integrate the intervention within the organization and to target behavior change in those who will deliver and support the intervention. A CFIR-based tool was created that can be adapted to assess a variety of health promotion programs targeting a variety of settings such as schools, child care, health services, and worksites. The inter-rater reliability results indicate that the code development process and coding process resulted in code definitions and coder scores that were of acceptable quality. This tool and methodology appear promising for identifying programs to prioritize for dissemination, guiding the selection of programs or program content to fit the needs of individual contexts, and identifying gaps across programs to support efforts to develop and test supplemental implementation supports.

Fifty-one CBPA-specific codes were developed to cover implementation factors from CFIR. The 23 codes for the Intervention Characteristics domain of CFIR were relatively straightforward to identify and assess. These codes generally captured the intervention options provided by the program and are likely considered critical factors when considering whether to adopt a program and which program to adopt. These characteristics transcend health promotion programs, and can be applied to a variety of interventions. A library or repository that provides information on these characteristics across available programs would provide a valuable resource from which end users could choose a program that aligns with their needs. As is known from various implementation theories including CFIR, information pertaining to Intervention Characteristics is necessary, but insufficient, for organizations to successfully implement programs. The 28 codes for the Inner Setting, Characteristics of Individuals, and Process domains of CFIR generally covered additional factors related to integrating the intervention within the organization, and promoting behavior change among those who have a role in intervention delivery. Implementation theory indicates that when this additional information is provided, programs are more successfully adopted, implemented, and sustained (Nilsen, 2015). However, the current study found that such information was often not included in existing programs for promoting CBPA. Thus, researchers and program developers should include and continue to test the value of reporting this information in programs.

In this study, some codes were challenging to use because they required reviewers to conduct detailed reviews of program materials and some subjectivity and judgment was required of the coder. Using a 0–2 scale (not covered, covered minimally, covered more than minimally) was advantageous because it allowed for the identification of content that addressed the code/factor more thoroughly, as minimal coverage is not likely to be of great benefit to supporting implementation. However, inter-rater reliability was less than ideal for some codes when using the 0–2 scale, suggesting that it was difficult for coders to identify the threshold used to represent more than minimal coverage. All codes had acceptable inter-rater reliability for the 0–1 scale, so this scale is preferred and the 0–2 scale should be used with caution.

Test-retest reliability values were acceptable for all indices. The advantage of the indices is to show the proportion of codes that were covered separately by domain to support inferences about overall programs. Programs that cover a large proportion of the codes could be prioritized for dissemination, and the related content from such programs could be used to supplement other programs (e.g., combining the best aspects of various programs).

Future studies are needed to continue to refine and expand this coding methodology to apply across health promotion programs. This tool and methodology appear promising for identifying programs to prioritize for dissemination, guiding the selection of programs or program content to fit the needs of individual contexts, and identifying gaps across programs to assist efforts to develop and test supplemental implementation supports.

4.1. Strengths and limitations

A strength of this work was the use of the CFIR, which is a comprehensive and commonly-used implementation science framework that has well-developed publicly available resources. One limitation of this study is that despite creating a 3-level scale to capture variability in the extent of meeting each code, few programs received a score of “more than minimally developed” on any of the codes. This may be due to the rarity of programs including exceptional implementation resources; however, another possibility is that the thresholds were poorly defined. Future studies should provide detailed definitions for coders to use in differentiating minimal versus extensive resources. Another limitation was that end user and stakeholder input was not directly solicited, other than reviewing what was covered in scientific publications. Authors did, however, communicate study findings with program sponsors (when contact information was available), in order to provide feedback that may be relevant to program revisions. Future studies should consider incorporating a participatory approach with end users to obtain perspectives about what users are seeking in health promotion programs and resources, although this likely varies by topic and content.

4.2. Conclusions

Implementation factors are critical determinants of implementation success (Tomoaia-Cotisel et al., 2013; Damschroder et al., 2009). One next step in supporting more widespread implementation and effectiveness of setting- and organization-based health promotion programs should be prioritizing the dissemination of programs that address theoretically-based implementation factors, and that supplement already-developed programs with implementation supports. This study describes a tool and methodology for assessing implementation factors within existing programs, which can be adapted for other health promotion programs to support their success and impact. Such tools and methodologies can be used by researchers and stakeholders to better

support practitioners and other end users to engage in successful implementation efforts. Practitioners can utilize this information to guide the selection of programs based on their individual needs and comprehensiveness of the program. A better consideration of real-world implementation factors in the development and dissemination of health promotion programs is likely needed for such programs to have an increased impact on public health.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pmedr.2019.100909.

Acknowledgements

Research reported in this publication was supported by the Katharine B. Richardson Research Award at Children's Mercy Kansas City, the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under Award Number F32DK115146, and the Institute of Education Sciences and U.S. Department of Education by grant R305A150277. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders.

Declaration of competing interest

The authors declare that there are no conflicts of interest.

References

Aarons, G.A., Hurlburt, M., Horwitz, S.M., 2011. Advancing a conceptual model of evidence-based practice implementation in public service sectors. Admin. Pol. Ment. Health 38 (1), 4–23 January 01.

Atkins, L., Francis, J., Islam, R., et al., 2017. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implement. Sci. 12 (1), 77 June 21.

Beets MW, Beighle A, Erwin HE, Huberty JL. After-school program impact on physical activity and fitness: a meta-analysis. Am. J. Prev. Med. Jun 2009;36(6):527–537.

Carlson, J.A., Sallis, J.F., Chriqui, J.F., Schneider, L., McDermid, L.C., Agron, P., 2013. State policies about physical activity minutes in physical education or during school. J. Sch. Health 83 (3), 150–156 Mar.

CFIR Research Team, 2015. CFIR Technical Assistance Website. Center for Clinical Management Research, Ann Arbor, MI Available at. http://cfirguide.org/constructs.html, Accessed date: 5 July 2018.

Cicchetti, D.V., 1994. Guidelines, criteria and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol. Assess. 6 (4).

Collins, J., Robin, L., Wooley, S., et al., 2002. Programs-that-work: CDC's guide to effective programs that reduce health-risk behavior of youth. J. Sch. Health 72 (3), 93–99 Mar.

Consolidated Framework for Implementation Research, 2014. CFIR Guide. http://www.cfirguide.org/cfirbib10.29.14.pdf, Accessed date: 1 December 2017.

Damschroder, L.J., Aron, D.C., Keith, R.E., Kirsh, S.R., Alexander, J.A., Lowery, J.C., 2009. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement. Sci. 4.

Durlak, J.A., DuPre, E.P., 2008. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. Am. J. Community Psychol. 41 (3–4), 327.

Glasgow, R.E., Emmons, K.M., 2007. How can we increase translation of research into practice? Types of evidence needed. Annu. Rev. Public Health 28 (1), 413–433.

Glasgow, R.E., Lichtenstein, E., Marcus, A.C., 2003. Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. Am. J. Public Health 93 (8), 1261–1267 2003/08/01.

Hoelscher, D.M., Springer, A.E., Ranjit, N., et al., 2010. Reductions in child obesity among disadvantaged school children with community involvement: the Travis County CATCH Trial. Obesity 18 (S1), S36–S44.

Kirk, M.A., Kelley, C., Yankey, N., Birken, S.A., Abadie, B., Damschroder, L., 2016. A systematic review of the use of the Consolidated Framework for Implementation Research. Implement. Sci. 11 (1), 1–13.

Kitson, A.L., Rycroft-Malone, J., Harvey, G., McCormack, B., Seers, K., Titchen, A., 2008. Evaluating the successful implementation of evidence into practice using the PARIHS framework: theoretical and practical challenges. Implement. Sci. 3.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33 (1), 159–174 Mar.

Nilsen, P., 2015. Making sense of implementation theories, models and frameworks. Implement. Sci. 10, 53 Apr 21.

Rabin, B.A., Brownson, R.C., Haire-Joshu, D., Kreuter, M.W., Weaver, N.L., 2008. A glossary for dissemination and implementation research in health. J. Public Health Manag. Pract. 14.

Sallis, J.F., McKenzie, T.L., Alcaraz, J.E., Kolody, B., Faucette, N., Hovell, M.F., 1997. The effects of a 2-year physical education program (SPARK) on physical activity and fitness in elementary school students. Sports, Play and Active Recreation for Kids. Am. J. Public Health 87 (8), 1328–1334 1997/08/01.

Sallis, J.F., Owen, N., Fisher, E., 2015. Ecological models of health behavior. In: Glanz, K., Rimer, B.K., Viswanath, K. (Eds.), Health Behavior: Theory, Research, and Practice, 5th edition. Jossey-Bass, San Francisco, CA, pp. 43–64.

Schwartz, M.B., Lund, A.E., Grow, H.M., et al., 2009. A comprehensive coding system to measure the quality of school wellness policies. J. Am. Diet. Assoc. 109 (7), 1256–1262 Jul.

Soler, R.E., Leeks, K.D., Razi, S., et al., 2010. A systematic review of selected interventions for worksite health promotion. Am. J. Prev. Med. 38 (2), S237–S262.

Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. Am. J. Prev. Med. Sep 2012a;43(3):337–350.

Tabak, R.G., Khoong, E.C., Chambers, D.A., Brownson, R.C., 2012b. Bridging research and practice: models for dissemination and implementation research. Am. J. Prev. Med. 43 (3), 337–350 9//.

Tomoaia-Cotisel, A., Scammon, D.L., Waitzman, N.J., et al., 2013. Context matters: the experience of 14 research teams in systematically reporting contextual factors important for practice change. Ann. Fam. Med. 11 (Suppl. 1), S115–S123 May-Jun.

Ward, D.S., Vaughn, A., McWilliams, C., Hales, D., 2010. Interventions for increasing physical activity at child care. Med. Sci. Sports Exerc. 42 (3), 526–534 2010/03//.