# Challenges in evaluating implementation and effectiveness in real-world settings: evaluation proposal for school-based health-promoting intervention

*Health Promotion International*, 2023, **38**, 1–7
https://doi.org/10.1093/heapro/daac185
**Perspectives**





Perspectives

# Challenges in evaluating implementation and effectiveness in real-world settings: evaluation proposal for school-based health-promoting intervention

**Marla T. H. Hahnraths[1,*], Maartje Willeboordse[2], and Onno C. P. van Schayck[1]**

[1]Department of Family Medicine, Care and Public Health Research Institute (CAPHRI), Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands
[2]Mulier Institute, P.O. Box 85445, 3508 AK Utrecht, The Netherlands
[*]Corresponding author. E-mail: mth.hahnraths@maastrichtuniversity.nl

## Summary

There are various research designs and approaches to investigate how health-promoting activities are implemented in complex, real-world systems, and to identify potential health effects that might occur following implementation. Although literature describes guidelines to perform and report about implementation research and effect evaluations, no specific guidelines exist on analysing and reporting about the combination of effectiveness data and implementation data collected as part of intervention evaluation in complex and diverse settings. This paper describes the evaluation of primary school-based health-promoting activities in complex systems. Furthermore, an approach for data categorization inspired by Rogers' Diffusion of Innovations theory is presented that can facilitate structuring the study's results and relating the degree of implementation to any impact on effectiveness outcomes that might be observed. Researchers interested in using this approach for data categorization have to ensure that the following three conditions are met: (i) data on an intervention's efficacy in a controlled setting with optimal implementation is available; (ii) key points that define an intervention's optimal implementation are available and (iii) an evaluation study is performed, collecting both effectiveness data and implementation data in a real-world context. This data categorization approach can be useful to generate more insight into an intervention's effectiveness under varying circumstances, and optimal support and advice can be provided to stakeholders to achieve maximum impact of population-based health-promoting interventions in complex, real-world systems. However, the proposed approach is a first suggestion and further testing and adaptation is necessary to increase its usefulness. Knowledge and experience sharing among researchers performing comparable research can increase the knowledge base regarding this subject.

**Keywords:** implementation, dissemination, complexity, research methodology

## BACKGROUND

For many decades, explanatory trials have generated an evidence base on the *efficacy* of numerous health-promoting interventions under controlled circumstances (Chau *et al.*, 2010; Sobol-Goldberg *et al.*, 2013; Wolfenden *et al.*, 2014). However, interventions' *effectiveness* in less controlled, complex, real-world systems often differs from their efficacy demonstrated in explanatory trials due to context- and implementation-related differences. Generating more knowledge on interventions' effectiveness in complex systems is therefore essential for successful dissemination. Health-promoting interventions are often implemented in complex systems consisting of multiple interacting components and characterized by adaptivity, non-linearity, feedback loops and the difficulty to control or predict outcomes (Patton, 2011; Moore et al., 2015, 2019; Rutter *et al.*, 2017). The complex-systems approach, which takes into account these characteristics, has been receiving increasing attention in the field of public health, with various researchers describing approaches

and frameworks for the evaluation of interventions in complex systems (Jolley, 2014; Rutter *et al.*, 2017; Moore *et al.*, 2019; Luna Pinzon *et al.*, 2022). Although taking a complex-systems approach (Rutter *et al.*, 2017) is one of the best ways to evaluate an intervention in a complex, real-world system, some of its characteristics (e.g. non-linearity) are not compatible with effectiveness and/or implementation evaluation. Furthermore, a complex-systems approach is a difficult and demanding approach, which is not always feasible to adopt. There is not always sufficient time, money and/or resources to investigate a specific system in the amount of detail that a complex-systems approach calls for. Various researchers have proposed different research designs and approaches to shed more light on the evaluation of an intervention in a real-world context. Schwartz and Lellouch described pragmatic trials, of which the primary aim is to determine an intervention's effect under *usual conditions* rather than under *ideal circumstances* (Schwartz and Lellouch, 1967; Thorpe *et al.*, 2009; Loudon *et al.*, 2015). There are several pragmatic trials evaluating the effects of health-promoting interventions in real-world settings, but the description of results is often limited to their impact on health outcomes and little implementation-related information is presented (Challen *et al.*, 2014; Long *et al.*, 2020). Effectiveness–implementation hybrid research designs combine elements from effectiveness and implementation research and therefore allow for the linking of outcomes from effectiveness studies to general implementation strategies and/or factors (Curran *et al.*, 2012). Curran *et al.* previously proposed three types of hybrid approaches: (i) testing intervention effects while gathering implementation-related information; (ii) dual testing of an intervention's effects and implementation strategies and (iii) testing an implementation strategy while gathering information regarding an intervention's effects on relevant outcomes (Curran *et al.*, 2012). Realist investigation is another often-described approach to evaluate an intervention's implementation and impact in complex settings. Using this approach, researchers aim to investigate what works for whom in which circumstances, instead of 'simply' investigating if a certain approach works (Tilley, 2000). Case studies are often mentioned as a suitable research design to answer realist investigations' research questions and to investigate an intervention's implementation and impact in complex and diverse settings, as they provide the opportunity to explore variance between different cases and to explain the circumstances under which certain outcomes are achieved (Woolcock, 2013; Pawson, 2019).

In this paper, we describe a multiple-case study evaluating various primary school-based health-promoting activities in complex, real-world systems and we reflect on various choices that have been made so far in this project. To inform and support other researchers and stakeholders in this field, we illustrate an approach for evaluating the implementation and effectiveness of interventions in complex, real-world systems and relating these findings to the results of previously executed explanatory trials on the efficacy of these interventions. Although we recognize the importance and value of the various principles of the complex-systems approach, we propose a simplified approach that can be of use in practice, when taking a complex-systems approach is not always feasible.

## RESEARCH EXAMPLE: THE HEALTHY PRIMARY SCHOOL OF THE FUTURE

The research that we will discuss as an example throughout this paper comprises the Healthy Primary School of the Future (HPSF), a school-based health-promoting intervention that was evaluated using a quasi-experimental design involving eight Dutch primary schools (four intervention schools and four control schools) (Willeboordse *et al.*, 2016). For 4 years, HPSF's efficacy on various outcomes was investigated, revealing significant positive intervention effects on outcomes such as children's body mass index *z*-score and dietary and physical activity (PA) behaviours (Bartelink *et al.*, 2019a,b; Willeboordse *et al.*, 2022). Following HPSF's efficacy, the ambition rose to spread its principles to other schools. However, this 'scaling-up' of the activity comes with several challenges. Schools can be defined as complex and unique systems that consist of various interacting components (e.g. teachers, children, parents, the school environment and the wider community) and have the ability to self-organize and adapt over time (Keshavarz *et al.*, 2010; Turunen *et al.*, 2017; Darlington *et al.*, 2018). Applying the complex-systems approach to school-based health promotion means acknowledging that what works in certain schools might not work in other schools. This underpins the importance of taking a school's context, population, wishes and needs into account when developing, implementing and evaluating health-promoting activities, as evidence suggests a strong interaction between an intervention and the context in which it is being implemented (Keshavarz *et al.*, 2010; Wang and Stewart, 2013; Turunen *et al.*, 2017; Darlington *et al.*, 2018; Kremers *et al.*, 2018). To generate knowledge on how HPSF-related activities are implemented in various real-world school contexts, to identify influential factors and to investigate the effectiveness on children's health and well-being, a follow-up research project of HPSF was initiated. In this project, 11 Dutch primary schools are followed between 2019 and 2023. The ambition at the start of the project was that all participating schools would eventually become full HPSFs. However, schools are free to decide if, when and to what degree they implement health-promoting activities in their setting, making the implementation

process less controlled than in the HPSF efficacy trial, in which intervention schools implemented a pre-defined intervention. Schools are responsible for their own development and implementation process, but they are aided by a process coordinator who works independently of the research team and supports each school during the project. Researchers play an observing role and have limited influence on the implementation processes in the various schools. This approach stimulates bottom-up development and implementation of pragmatic and school-specific activities, which is hypothesized to facilitate sustainable integration of health within the schools.

The project's real-world nature provides several challenges when it comes to choosing a suitable research design that can account for the complexity of the setting. All schools are working on integrating health within their organization, meaning that no control group is included in the project. Furthermore, instead of a pre-defined, standardized intervention, there is great variety in activities implemented in the schools, making often-utilized implementation outcomes such as fidelity (the degree to which an intervention is delivered as intended) less relevant in the present study. Rather, there is more emphasis on the actual implementation (*what* is implemented in the various schools?) and the reasons behind this implementation (*why* is implementation (not) successful?). Compared with the HPSF efficacy trial, which was mainly explanatory, the follow-up research project is therefore more pragmatic in nature and resembles a realist evaluation (Schwartz and Lellouch, 1967; Pawson and Tilley, 1997; Thorpe *et al.*, 2009; Loudon *et al.*, 2015). The aim of this research is to investigate the implementation of HPSF-related activities under complex, real-world circumstances and to explore the potential influence of differences in implementation on HPSF's effectiveness on various health outcomes. The lessons learned from this research can then be used to disseminate HPSF's principles to other schools. The schools' pre-existing contexts and the 'natural' implementation process of HPSF-related activities in a real-world context are investigated by observing relevant meetings with stakeholders (e.g. directors, managers and teachers), gathering data via questionnaires and performing semi-structured interviews with various stakeholders (directors, managers and teachers). Besides data on actual implementation (*what* is implemented in the various schools?), information on reasons behind this implementation and potential implementation differences between schools (*why* is implementation in school A successful and *why* is implementation not successful in school B?) is gathered through these sources. All implementation data are structured using the characteristics from the framework by Fleuren, Wiefferink and Paulussen (characteristics of the socio-political context, organization, person, innovation and innovation strategy) (Fleuren *et al.*, 2004). For example, if one school is successful in implementing an extensive activity (e.g. the provision of a daily healthy school lunch) and interview participants from this school mention the positive influence of working on this implementation with the complete team while another school is not successful in implementing a daily school lunch as the activity is only carried by the school director, these differences in implementation can (partly) be explained by differences in characteristics of the innovation strategy. Besides data on the implementation process, the same data on children's anthropometrics, dietary and PA behaviours and well-being are gathered as in the HPSF efficacy trial to compare the impact on these outcomes in both settings. This will provide an estimate of the effectiveness of HPSF-related activities under real-world conditions compared with their efficacy following maximum implementation under controlled circumstances.

## Challenges in data categorization, analysis and reporting

Although literature describes various guidelines to perform and report about implementation research and effect evaluations, no specific guidelines or practical approaches exist on how to analyse and report about the combination of effectiveness data and implementation data collected in real-world and complex settings (Moore *et al.*, 2015; Neta *et al.*, 2015). To deal with this and to explore if differences in implementation lead to differences in HPSF's impact on children's health and well-being, a novel approach for data categorization, analysis and reporting is proposed. This approach is inspired by Rogers' Diffusion of Innovations theory, which states that a population can be divided into five adopter categories (innovators, early adopters, early majority, late majority and laggards) based on their degree of innovativeness regarding an introduced innovation. Subjects with a high degree of innovativeness (e.g. innovators or early adopters) will adopt an introduced innovation faster than those with a lower degree of innovativeness (e.g. late majority or laggards) (Rogers, 2010). In the present project, schools are not introduced to a standardized innovation. Rather, they have a shared ambition (integrating health within their organization and—if realistic—becoming full HPSFs) and consequently shape activities fitting their context throughout the project. These activities can be composed of different components with varying intensity (e.g. changes in a school's policy, practices and/or communication strategies) that evolve over time. To be able to compare participating schools with each other and with the efficacy trial's full HPSFs, they will be divided over Rogers' adopter categories based on their degree of innovativeness. This degree of innovativeness is operationalized through the number of HPSF key points that schools adhere to 3 years after the

start of the project. A period of 3 years was chosen as during the HPSF efficacy trial, it was observed that it took several years before project-specific plans and ambitions were formulated and implementation started. Additionally, the COVID-19 pandemic greatly restricted schools' ability to work on the present project, thereby delaying project development and implementation. Four HPSF key points defining optimal implementation of HPSF were formulated by stakeholders involved in the efficacy trial to facilitate further dissemination of HPSF. For optimal implementation, various stakeholders (e.g. school staff, children and parents) should be actively involved (key point 1) and a school-wide approach should be taken, meaning that implemented activities should reach all children within a school (key point 2). In the HPSF efficacy trial, these two key points were perceived as vital to achieve successful implementation of HPSF. Furthermore, optimal implementation of HPSF means that children engage in at least one hour of PA every day (key point 3) and consume a daily healthy lunch at school (key point 4). For a detailed description of the HPSF key points, see Additional File 1. If a school participating in the present research project adheres to all four HPSF key points at the end of data collection (i.e. the school can be considered a full HPSF), this is defined as having the highest degree of innovativeness and the school will be categorized in the innovators/early adopters category (these two adopter categories are combined for clarity reasons). Alternatively, schools who do not adhere to any key points at the end of data collection will have the lowest degree of innovativeness and consequently fall in the laggards category. A further specification of the categorization of schools based on HPSF key points can be found in Table 1. Although the term degree of innovativeness might imply that we assume that differences in actual implementation between the various schools are simply the result of differences in degree of innovativeness, we acknowledge these differences can be caused by various factors that are much more diverse than differences in degree of innovativeness only. The term degree of innovativeness is solely used for clarity reasons and should not be used to explain implementation differences as it disregards the diversity and complexity of various real-world settings. Instead, the implementation processes of HPSF in the various schools are continuously investigated through various methods (e.g. observing relevant meetings with stakeholders, gathering data via questionnaires and performing semi-structured interviews with various stakeholders) to obtain a more extensive impression of (reasons behind) implementation differences between the various schools, although this goes beyond the scope of the present paper and will therefore not be further discussed here. The described approach for categorization will ease comparison between the various schools participating in the present study as it provides a potential framework to structure the study's results. When analysing and reporting about the study's implementation and/or impact on health outcomes, the data and an article's results section can be structured using the different adopter categories instead of the more traditional intervention

**Table 1:** Specification of school categorization based on HPSF key points

| | HPSF key points | | | |
|---|---|---|---|---|
| | **Actively involved stakeholders (e.g. parents, children, school staff)** | **School-wide approach** | **Children engage in at least 1 h of PA per day** | **Children consume a daily healthy lunch** |
| Original efficacy trial (optimal implementation of HPSF) | | | | |
| Full HPSFs | X | X | X | X |
| Follow-up research project (scaled-up, real-world context) | | | | |
| Adopter categories | | | | |
| Innovators/early adopters | X | X | X | X |
| Early majority | X | X | X[a] | X[a] |
| Late majority | X | | | |
| Laggards | | | | |

*Note*. A more elaborate description of the four HPSF key points can be found in Additional File 1. Key points 1 and 2 are assumed to be necessary for schools to achieve key points 3 and 4, which is why there is no categorization for a combination of key point 1 or 2 with key point 3 and/or 4. In the follow-up research project, the categorization of schools will be done at one moment in time, i.e. 3 years after the start of the project (at the end of data collection). This means that the categorization will be based on a snapshot of a longitudinal, ongoing implementation process. Categorization at a different moment in time could therefore lead to different results.
[a]Schools have to adhere to one of these two indicated key points to fall in the early majority category.

and control group i.e. often used to describe an explanatory trial's results. The approach however is solely a first suggestion and further refinements and improvements are strongly advised to increase its usefulness and validity. For example, the categorization of the schools in the different adopter categories is rather robust when using dichotomized key points as proposed in this paper (i.e. categorizing schools based on whether or not they adhere to the various key points). This limits the sensitivity with which the implementation degree of the activities can be related to any impact on children's health and well-being that might be observed, as not all differences in schools' implementation degrees will be captured and acknowledged using this approach. Rather than simply observing if schools do or do not adhere to the key points at the end of data collection, the key points could be addressed in a more continuous matter to obtain a more sensitive degree of implementation. For example, a school providing a healthy school lunch on a daily basis could be assigned a higher implementation degree than a school providing a healthy school lunch three times a week. By acknowledging these more subtle differences, the refined approach would provide a more nuanced and sensitive degree of implementation for the various schools, which can subsequently be linked to any impact on health outcomes that might be observed. Following this approach could make it possible to provide more general recommendations regarding HPSF implementation in complex systems. The intervention effects observed in the previously executed HPSF efficacy trial are assumed maximal due to the controlled conditions under which the complete intervention was implemented. By using the refined categorization approach and comparing the results from the efficacy trial with the observations in various real-world settings, it would be possible to identify certain elements of HPSF that lead to more health-related impact than others, and to recognize general implementation factors and/or strategies that would facilitate the implementation of these elements in various settings.

Besides the limitations related to the current robustness of data categorization, it should be noted that categorization with this approach happens at one moment in time, and categorization at a different moment could therefore lead to different results. This is in line with the continuously changing complex and real-world conditions and settings in which the research is executed. Implementation is thought to be a continuous process i.e. never finished and to the best of our knowledge, there is currently no model or categorization approach available that accounts for the continuously shifting degree of implementation over time.

## IMPLICATIONS

This novel approach for data categorization can be useful to other scientists performing comparable research on the implementation and impact of general activities in a complex, real-world system following an efficacy trial with an observed promising impact on relevant outcomes. This research does not have to be limited to health promotion in the school setting but can span a wide range of research areas. The additional benefit of this approach is that an intervention's relative effectiveness on relevant outcomes following a certain degree of implementation can be established when compared with the effects observed in a previously performed efficacy trial. As a result, more insight into what works under which circumstances is generated and optimal support and advice can be provided to stakeholders to achieve maximum impact of population-based health-promoting interventions in complex, real-world systems. It should however be taken into account that categorization with this approach is done at one specific moment in time, and categorization at a different moment could lead to different results due to the complexity and adaptivity of the context and the implementation process that will continue to develop. Furthermore, categorization as proposed in its current, dichotomous form is rather robust and further refinement is strongly recommended to increase the approach's usefulness and potential. The approach is not yet empirically tested and elaboration and adaptation are therefore necessary for further improvement and to increase its validity. As previously mentioned by Huiberts *et al.* (Huiberts *et al.*, 2022), researchers sharing their experiences, insights and approaches regarding evaluating interventions in complex, real-world systems can greatly facilitate the development of adequate and feasible evaluation approaches and should therefore be encouraged. Researchers interested in using the approach for data categorization proposed in this paper have to ensure that the following three conditions are met: (i) data on an intervention's efficacy in a controlled setting with optimal implementation are available; (ii) key points that define an intervention's optimal implementation are available and (iii) an evaluation study is performed, collecting both effectiveness data and implementation data in a real-world setting.

## CONCLUSION

To advance research and to stimulate intervention dissemination and sustainability, it is vital to investigate how (preventive) activities are implemented in complex and real-world systems, and to identify potential health effects that might occur following this implementation. By combining elements from efficacy, effectiveness and implementation research, outcomes can be related to general implementation strategies and/or factors. The

approach for data categorization described in this paper can be useful to relate an intervention's impact to specific implementation strategies and/or factors and thereby generating more insight into what works under which circumstances. This can subsequently lead to improved support and advice provision to stakeholders aiming to achieve maximum impact of population-based (health-promoting) interventions in complex, real-world systems. We acknowledge that the approach described in this paper describes a simplified version of reality and does not take into account all principles of a systems approach (e.g. non-linearity). Nonetheless, it could be of benefit in implementation science, where taking the complete complex-systems approach is not always feasible. However, further testing, adaptation and refinement of the approach are necessary to increase its usefulness and validity. Knowledge and experience sharing among researchers working on comparable issues can increase the knowledge base regarding evaluating interventions in complex, real-world systems.

## Supplementary Material

Supplementary material is available at *Health Promotion International* online.

## Funding

This work was supported by the Limburg provincial authorities (www.limburg.nl) and Maastricht University [project number SAS-2019-00624]. The funder had no role in the design, execution, interpretation or writing of the study.

## Authors' Contributions

M.T.H.H.: conceptualization, methodology, writing—original draft; M.W. and O.C.P.v.S.: conceptualization, methodology, writing—review and editing.

## Conflict of Interest

The authors declare that they have no competing interests.

## Ethical Approval and Consent to Participate

Not applicable.

## Consent for Publication

Not applicable.

## Data Availability

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

## REFERENCES

Bartelink, N. H. M., van Assema, P., Kremers, S. P. J., Savelberg, H. H. C. M., Oosterhoff, M., Willeboordse, M. *et al.* (2019a) Can the healthy primary school of the future offer perspective in the on-going obesity epidemic in young children? A quasi-experimental study. *BMJ Open*, **9**. doi:10.1136/bmjopen-2019-030676

Bartelink, N. H. M., van Assema, P., Kremers, S. P. J., Savelberg, H. H. C. M., Oosterhoff, M., Willeboordse, M. *et al.* (2019b) One- and two-year effects of the healthy primary school of the future on children's dietary and physical activity behaviours: a quasi-experimental study. *Nutrients*, **11**, 1–18.

Challen, A. R., Machin, S. J. and Gillham, J. E. (2014) The UK Resilience Programme: a school-based universal nonrandomized pragmatic controlled trial. *Journal of Consulting and Clinical Psychology*, **82**, 75–89.

Chau, J. Y., Van Der Ploeg, H. P., van Uffelen, J. G. Z., Wong, J., Riphagen, I., Healy, G. N. *et al.* (2010) Are workplace interventions to reduce sitting effective? A systematic review. *Preventive Medicine*, **51**, 352–356.

Curran, G. M., Bauer, M., Mittman, B., Pyne, J. M. and Stetler, C. (2012) Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Medical Care*, **50**, 217–226.

Darlington, E. J., Violon, N. and Jourdan, D. (2018) Implementation of health promotion programmes in schools: an approach to understand the influence of contextual factors on the process? *BMC Public Health*, **18**. doi:10.1186/s12889-017-5011-3

Fleuren, M., Wiefferink, K. and Paulussen, T. (2004) Determinants of innovation within health care organizations: literature review and Delphi study. *International Journal for Quality in Health Care*, **16**, 107–123.

Huiberts, I., Singh, A., van Lenthe, F. J., Chinapaw, M. and Collard, D. (2022) Evaluation proposal of a national community-based obesity prevention programme: a novel approach considering the complexity perspective. *International Journal of Behavioral Nutrition and Physical Activity*, **19**, 1–12.

Jolley, G. (2014) Evaluating complex community-based health promotion: Addressing the challenges. *Evaluation and Program Planning*, **45**, 71–81. doi:10.1016/J.EVALPROGPLAN.2014.03.006

Keshavarz, N. M., Nutbeam, D., Rowling, L. and Khavarpour, F. (2010) Schools as social complex adaptive systems: a new way to understand the challenges of introducing the health promoting schools concept. *Social Science and Medicine*, **70**, 1467–1474.

Kremers, S. P. J., Visscher, T. L. S. and Schuit, A. J. (2018) Effect in zijn context [Effect in its context]. *Tijdschrift Voor Gezondheidswetenschappen*, **96**, 128–131.

Long, M. W., Weber, M. R., Allan, M. J., Ma, Y., Jin, Y., Aldous, A. *et al.* (2020) Evaluation of a pragmatic trial of a collaborative school-based obesity prevention intervention in a low-income urban district. *Preventive Medicine*, **133**, 106020.

Loudon, K., Treweek, S., Sullivan, F., Donnan, P., Thorpe, K. E. and Zwarenstein, M. (2015) The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ*, **350**, h2147.

Luna Pinzon, A., Stronks, K., Dijkstra, C., Renders, C., Altenburg, T., den Hertog, K. *et al.* (2022) The ENCOMPASS framework: a practical guide for the evaluation of public health programmes in complex adaptive systems. *International Journal of Behavioral Nutrition and Physical Activity*, **19**, 1–17. doi:10.1186/S12966-022-01267-3

Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W. *et al.* (2015) Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, **350**. doi:10.2307/26520417

Moore, G. F., Evans, R. E., Hawkins, J., Littlecott, H., Melendez-Torres, G. J., Bonell, C. *et al.* (2019) From complex social interventions to interventions in complex social systems: Future directions and unresolved questions for intervention development and evaluation. *Evaluation*, **25**, 23–45. doi:10.1177/1356389018803219

Neta, G., Glasgow, R. E., Carpenter, C. R., Grimshaw, J. M., Rabin, B. A., Fernandez, M. E. *et al.* (2015) A framework for enhancing the value of research for dissemination and implementation. *American Journal of Public Health*, **105**, 49–57.

Patton, M. (2011) *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. The Guilford Press, New York, NY, USA.

Pawson, R. (2019) Pragmatic trials and implementation science: grounds for divorce? *BMC Medical Research Methodology*, **19**, 1–13.

Pawson, R. and Tilley, N. (1997) An introduction to scientific realist evaluation. In Chelimsky, E. and Shadish, W. R. (eds), *Evaluation for the 21st Century: A Handbook*. SAGE Publications, Inc., Thousand Oaks, CA, USA, pp. 405–418. doi:10.4135/9781483348896

Rogers, E. M. (2010) *Diffusion of Innovations*, 4th edition. Simon & Schuster Inc, New York, NY, USA.

Rutter, H., Savona, N., Glonti, K., Bibby, J., Cummins, S., Finegood, D. T. *et al.* (2017) The need for a complex systems model of evidence for public health. *The Lancet*, **390**, 2602–2604. doi:10.1016/S0140-6736(17)31267-9

Schwartz, D. and Lellouch, J. (1967) Explanatory and pragmatic attitudes in therapeutical trials. *Journal of Chronic Diseases*, **20**, 637–648.

Sobol-Goldberg, S., Rabinowitz, J. and Gross, R. (2013) School-based obesity prevention programs: a meta-analysis of randomized controlled trials. *Obesity*, **21**, 2422–2428.

Thorpe, K. E., Zwarenstein, M., Oxman, A. D., Treweek, S., Furberg, C. D., Altman, D. G. *et al.* (2009) A pragmatic–explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *Journal of Clinical Epidemiology*, **62**, 464–475.

Tilley, N. (2000) Realistic evaluation: an overview. Founding Conference of the Danish Evaluation Society. https://citeseerx.ist.psu.edu/viewdoc/download?-doi=10.1.1.1058.7583&rep=rep1&type=pdf (last accessed 15 May 2022).

Turunen, H., Sormunen, M., Jourdan, D., Von Seelen, J. and Buijs, G. (2017) Health promoting schools—a complex approach and a major means to health improvement. *Health Promotion International*, **32**, 177–184.

Wang, D. and Stewart, D. (2013) The implementation and effectiveness of school-based nutrition promotion programmes using a health-promoting schools approach: a systematic review. *Public Health Nutrition*, **16**, 1082–1100.

Willeboordse, M., Bartelink, N. H. M., van Assema, P., Kremers, S. P. J., Savelberg, H. H. C. M., Hahnraths, M. T. H. *et al.* (2022) Battling the obesity epidemic with a school-based intervention: long-term effects of a quasi-experimental study. *PLoS One*, **17**, e0272291. doi:10.1371/JOURNAL.PONE.0272291

Willeboordse, M., Jansen, M. W., van den Heijkant, S. N., Simons, A., Winkens, B., de Groot, R. H. M. *et al.* (2016) The healthy primary school of the future: study protocol of a quasi-experimental study. *BMC Public Health*, **16**, 639.

Wolfenden, L., Wyse, R., Nichols, M., Allender, S., Millar, L. and McElduff, P. (2014) A systematic review and meta-analysis of whole of community interventions to prevent excessive population weight gain. *Preventive Medicine*, **62**, 193–200.

Woolcock, M. (2013) Using case studies to explore the external validity of 'complex' development interventions. *Evaluation*, **19**, 229–248.