

The process evaluation of a school-based physical activity intervention: influencing factors and potential consequences of implementation

School-based
physical
activity
intervention

121

Received 12 January 2020
Revised 1 April 2020
Accepted 21 April 2020

Andreas Åvitsland, Stein Erik Ohna and Sindre Mikal Dyrstad
University of Stavanger, Stavanger, Norway
Hege Eikeland Tjomsland and Øystein Lerum
Western Norway University of Applied Sciences, Bergen, Norway, and
Eva Leibinger
University of Stavanger, Stavanger, Norway

Abstract

Purpose – This paper evaluates the implementation of a school-based physical activity intervention and discusses how the intervention outcomes can be influenced by the implementation.

Design/methodology/approach – In four of the nine lower secondary schools in which the intervention was conducted, the authors examined implementation fidelity, adaptation, quality, responsiveness and dose received. The authors conducted focus group interviews with teachers ($n = 8$) and students ($n = 46$) and made observations. Dose delivered was examined quantitatively, with weekly registrations.

Findings – Results showed that two out of four schools made few and positive adaptations, implemented the intervention with high fidelity and quality and responded positively. Four main factors were found to influence implementation: frame factors, intervention characteristics, participant characteristics and provider characteristics.

Research limitations/implications – A cross-sectional design was used and may not represent implementation throughout the whole school year.

Practical implications – In terms of large-scale implementation, the intervention may be generalizable. However, intervention criteria such as adequate facilities and a flexible timetable may be unattainable for some schools. The intervention can be adapted without compromising its purpose, but adaptations should be a result of cooperation between students and teachers.

Originality/value – Process evaluations on this topic are rare. This study adds to a limited knowledge base concerning what factors may influence implementation of school-based physical activity interventions for adolescents.

Keywords Process evaluation, School, Intervention, Physical activity, Implementation, Qualitative methods

Paper type Research paper

Introduction

Schools are considered viable settings for intervention, reaching children and adolescents irrespective of sex, socioeconomic status and ethnicity. Studies evaluating school-based physical activity (PA) interventions have largely examined whether the interventions affect

©Andreas Åvitsland, Stein Erik Ohna, Sindre Mikal Dyrstad, Hege Eikeland Tjomsland, Øystein Lerum and Eva Leibinger. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>.



Health Education
Vol. 120 No. 2, 2020
pp. 121-139
Emerald Publishing Limited
0965-4283
DOI 10.1108/HE-01-2020-0004

students' academic achievement (Norris *et al.*, 2015), physical activity and fitness (Kriemler *et al.*, 2011) or mental health (Smith *et al.*, 2018; Lubans *et al.*, 2016). The way in which such interventions are implemented may partly determine their effectiveness (Durlak and Dupre, 2008). However, monitoring and evaluating implementations has not been prioritized in school-based PA intervention research (Naylor *et al.*, 2015; Watson *et al.*, 2017; Daly-Smith *et al.*, 2018), increasing the risk of a type 3 error (Dobson and Cook, 1980). A type 3 error can occur when researchers dismiss a potentially effective intervention based on unsatisfactory results, when, in fact, the results were caused by poor implementation and not the intervention itself. Among the existing studies evaluating school-based PA interventions, only a few have addressed the link between implementation and outcome (Naylor *et al.*, 2015), thus limiting our understanding of their relationship.

Recent reviews of school-based PA interventions have shown mixed results in terms of physical, mental and cognitive outcomes (Demetriou and Honer, 2012; Daly-Smith *et al.*, 2018; Watson *et al.*, 2017; Singh *et al.*, 2018; Hynynen *et al.*, 2016). One reason for this may have to do with the difference between efficacy and effectiveness. Efficacy refers to an intervention's effect under ideal circumstances, while effectiveness refers to an intervention's effect under real-life conditions (Revicki and Frank, 1999). Conducting interventions in real-life conditions, such as in schools, makes it more difficult to assess effects, since variables outside the researcher's control may come into play. Additionally, schools are complex social systems (Moore *et al.*, 2019), because the individual agents (e.g. students, teachers) within the system are "numerous, dynamic, autonomous, highly interactive, learning and adaptive" (Keshavarz *et al.*, 2010). These agents and other school factors can influence implementation (Clarke, 2010). For instance, in their review of school-based PA interventions, Naylor *et al.* (2015) identified 20 factors that could hamper implementation. "Limited time" was the most frequently mentioned factor, followed by "lack of resources," "lesson scheduling" and "weather." The outcomes from a school-based cluster randomized controlled trial (RCT) may, therefore, have more to do with the specific school context than the intervention itself.

Interventions can also be complex, posing further challenges to evaluating outcomes (Craig *et al.*, 2008). When a complex intervention is introduced in several unique complex contexts, dynamic processes and interactions can lead to the intervention evolving differently in the different contexts. This could in turn influence the intervention outcomes (Hawe *et al.*, 2009) and obstruct researchers' ability to make causal inferences (Rickles, 2009). However, by monitoring and evaluating the process of implementation, we can improve our ability to interpret outcomes and causes (Oakley *et al.*, 2006; Tarp *et al.*, 2016).

This paper reports on the process evaluation of the "Don't worry, be happy" (DWBH) intervention, one of two separate intervention arms in a cluster RCT called School in Motion (ScIM). ScIM was designed to assess whether 120 min of additional weekly school time PA affected students' academic achievement, learning environment, physical fitness, PA levels and mental health. The findings of the process evaluation should be used as a supplementary tool when interpreting the outcomes of the ScIM study, which will be published in upcoming papers.

Aim and research questions

The aim of this research was to evaluate the implementation of the DWBH intervention. Furthermore, in the context of searching for feasible methods to increase school time PA for adolescents, we aim to increase our understanding of what works under what circumstances and why. In order to reach these aims, we asked the following questions:

- (1) How was the intervention implemented?
- (2) What influenced the implementation?

Theoretical background

The purpose of a process evaluation is to “explore the implementation, receipt, and setting of an intervention and help in the interpretation of the outcome results” (Oakley *et al.*, 2006). The DWBH intervention can be seen as a complex intervention, which according to Craig *et al.* (2008), is characterized by (1) the number of interacting components, (2) the number and difficulty of behaviors required by those delivering or receiving the intervention, (3) the number of groups or organizational levels targeted by the intervention, (4) the number and variability of outcomes and (5) the degree of flexibility. The present study understands implementation as defined by Durlak (2016): “. . .the ways a programme is put into practice and delivered to participants.” According to Durlak and Dupre (2008), implementation consists of eight separate but overlapping aspects. The present study addresses five: fidelity, adaptation, quality, responsiveness and dose delivered. Dose received does not fall under implementation as defined by Durlak (2016) but has been included as a sixth aspect in this evaluation to help understand how relevant the intervention is to the target participants (Linnan and Steckler, 2002). While dose delivered refers to the specific amount of intervention that has been provided to its participants, dose received refers to actual participation or attendance (Berkel *et al.*, 2011). Including both can offer important information about their relationship.

Besides dose received, focusing on the aforementioned aspects draws on the argumentation by Berkel *et al.* (2011), that they “occur within the delivery of program sessions, and as a result, constitute potential sources of disconnect between the program as designed and that which is implemented”. While fidelity, adaptation, quality and dose delivered are mainly determined by the intervention providers, responsiveness and dose received are determined by the participants, allowing us to examine the dynamic relationship between provider and participant, who can influence each other and ultimately, the intervention results (Berkel *et al.*, 2011). Evaluating quality has been neglected in previous implementation evaluations of school-based PA interventions (Naylor *et al.*, 2015). Humphrey *et al.* (2017) argue, however, that quality may be more important than fidelity and dose when it comes to impacting study outcomes; so this aspect was included. The other three aspects of implementation – differentiation, monitoring control and reach – can also influence the outcomes, but are not explicitly part of the program delivery or the relationship between provider and participant. We therefore chose not to focus on these aspects.

Methods

The School in Motion study

The ScIM study was a multicenter trial conducted in Norway during the 2017–2018 school year. The study was initiated by the Norwegian Directorate of Health and the Directorate of Education and Training. The interventions were designed by the project management group at the Norwegian School of Sports Sciences, piloted with a small number of classes and adjusted, before commencing the ScIM study. The project management group invited 103 lower secondary schools to participate, and 29 schools accepted. Only students attending ninth grade (14–15 years) during the intervention period were included. The intervention period lasted 29 weeks and since the intervention would be part of the school’s curriculum, all the 2,733 eligible students were required to participate, as they would in any other school subject. The DWBH intervention was assigned to nine schools (663 eligible students). Each school chose one teacher liaison, who was responsible for the intervention execution and for communicating with the researchers at their respective test center. The ScIM study was approved by the Norwegian Centre for Research Data (project number 49094).

The Don’t worry, be happy intervention. The goal of the DWBH intervention is to facilitate positive experiences with PA in order to contribute to healthy adolescent development.

A summary of key DWBH intervention characteristics and their relationship with fidelity, adaptation, quality, responsiveness and dose delivered/received are presented in [Table 1](#). DWBHs theoretical framework and components are extensively described elsewhere ([Kolle et al., 2020](#)). Briefly, DWBH draws on the theoretical perspectives on positive youth development ([Lerner, 2015](#)), relational developmental systems theory ([Lerner et al., 2015](#)) and positive movement experiences ([Agans et al., 2013](#)). According to the intervention’s underlying theories, the goal of the DWBH intervention is achieved by allowing the students individuality and personal interest in their choice of activities. Furthermore, the activities need to take place in a social environment that enables the participants to experience competence, confidence, connection, character and caring.

The intervention consisted of two separate weekly lessons, “Don’t worry” (DW) and “Be happy” (BH). DW could be conducted as a regular physical education (PE) lesson, although we encouraged the teachers to allow students to pursue activities of their choice. BH was the main component and requires a more detailed description. There was an initial planning phase in which students were asked to choose an activity or sport they wanted to pursue for the rest of the school year. Next, students who chose the same activity formed groups and started planning long-term activity goals, conflict solution strategies and organized a leadership structure. The plans were formalized in a written “activity contract” (see attachment 1), which was signed by the group members and approved by the teacher. Once approved, the planning phase was over, and the students pursued their activities in the BH lessons for the rest of the school year. Students were not assessed in BH. Although BH was student-led, teachers were the formal providers of the intervention and were required to be qualified PE teachers. Their tasks were to be present, observe and provide guidance when necessary. BH was to be organized so that students could participate together across homeroom classes.

Guidelines were provided for the incorporation of the two DWBH lessons into the school’s schedule. To schedule one of the lessons, the schools were required to reallocate 5% of the time from other subjects. The other lesson was added on top of the existing schedule. This

Table 1.
Summary of key
intervention
characteristics and
their relationship with
the implementation
aspects

Fidelity	(1) “Don’t worry”: Similar to ordinary PE in separate classes but students have the opportunity to choose their activity. “Be happy”: students pursue activities of their choice, in groups across homeroom classes, that they formed themselves (2) Students choose activity based on interest, not based on who they can be together within the group. Groups should stay together. Maximum eight students per group (3) Sufficient information must be given to the students. Students must use the activity contract to conduct long-term planning. All groups must have a leadership structure, group goals and plan for conflict resolution (4) 5% of time from other subjects should be taken to make room for one lesson. The second comes in addition to the ordinary schedule
Quality	(1) Sufficient facilities and equipment (2) Sufficient teacher-to-student ratio (3) Teachers should be available and able to help when necessary. They should interfere when necessary by recognizing the need for flexibility, evaluation, group alterations and conflict resolution
Responsiveness	(1) Positive response toward DWBH, and regarding the intervention as relevant, useful, advantageous in any way (2) Responsiveness also applies to how the teachers respond to the intervention
Dose delivered/ received	(1) Two lessons per week during the 29-week intervention period
Adaptations	(1) No predetermined adaptations have been defined. Small adaptations can be made if necessary and/or if they benefit overall implementation

resulted in students having a 45–60 min longer school day once a week, for which the schools were economically compensated. The schools were free to choose when to conduct the two lessons but were encouraged to choose two separate days of the week.

The process evaluation

Design and participants. This process evaluation used a cross-sectional design to gather qualitative data by conducting interviews and observations. A longitudinal design was used to collect quantitative data, by teachers reporting dose delivered each week throughout the intervention period. Participant sampling for the interviews was a combination of random and purposive. We randomly selected four schools assigned DWBH, one from each of the four regions in Norway where the intervention was carried out. School 1 included 52 students from two classes and was located in a rural area outside a major city. School 2 included 117 students from four classes and was located in a residential/rural area outside a major city. School 3 included 47 students from two classes and was located in a residential/urban area close to a smaller city. School 4 included 87 students from four classes and was located in a residential/urban area between two moderate-sized cities. Teacher liaisons at each school accepted the invitation to participate in the process evaluation. Next, the teacher liaisons were asked to perform a purposive sampling of students to be interviewed: three activity groups representing different activities and opinions toward the intervention and PA in general. The purposive sampling strategy was employed to cover a diversity of experience, in order to prevent bias toward presenting only one type of information from one type of participants (Robinson, 2014). Teachers who supervised DWBH were also interviewed to obtain knowledge about the implementation process from the providers' perspectives. A total of 54 individuals were interviewed. This amounted to 12 student focus group interviews ($n = 46$), two individual teacher interviews and two teacher focus group interviews ($n = 6$). Student interviewees provided written informed consent from their parents, and teacher interviewees gave their consent verbally. This study has been designed, conducted and reported so as to ensure the confidentiality and anonymity of participants.

Data collection. We conducted semistructured focus group interviews and individual interviews to capture participants' and providers' experiences of the intervention. Initially, all the interviews were supposed to be conducted in groups, but two teacher interviews were individual because of illness among the teachers. The interview guide was constructed to elicit answers that could be linked to the six included implementation aspects (see attachment 2). We chose semistructured interviews based on our expectation of broad variation of opinions and experiences regarding DWBH and other emerging issues (Barriball and While, 1994). Focus group interviews are suitable for program evaluation because participants answer questions about a specific topic in a social context; they can discuss and potentially reveal information that would not have emerged in an individual interview (Frey and Fontana, 1991).

Secondary data were gathered (Manzano, 2016) by observing one BH lesson in each of the selected schools. Observations were conducted to experience a physical presence, which provides an impression of the environmental surroundings, the participants, attitude toward the intervention and the dynamics between provider and participants. The purpose of these impressions was to assist in analyzing the interviews.

Qualitative data were gathered within the same week in each school, during the second half of the intervention period (between mid-January and mid-March 2018). The interviews took place during school hours at the participants' respective schools, in a classroom with only the researcher and the interviewee(s) present. The interviews lasted between 30 and 55 min and were audio recorded.

We defined dose delivered as the percentage of lessons provided to the participants, relative to the number of lessons that were possible to provide during the intervention period.

To measure dose delivered, the teacher liaisons used an online registration tool to weekly register DWBH lessons as executed/not executed. We considered 80% and above to be a high delivered dose. It is important to note that dose received is also a quantitative concept and by assessing it qualitatively, we only get an indication of how attendance has been relative to the dose delivered.

Data analysis. Audio recordings of the interviews were imported into NVivo qualitative data analysis Software 12 (QSR International Pty Ltd., Doncaster, Australia) and transcribed verbatim. Data were further analyzed in NVivo and Excel, using the five steps of the framework analysis (Spencer and Ritchie, 2002). The rationale behind choosing this approach is that it exists within the family of content or thematic analysis and can be used to “identify commonalities and differences (. . .) focusing on relationships between different parts of the data, thereby seeking to draw descriptive and/or explanatory conclusions” (Gale *et al.*, 2013). How we followed the five steps is outlined in Figure 1. Briefly, one deductive analysis and one inductive analysis were conducted consecutively. The deductive analysis was guided by research question 1 (how the intervention was implemented), while the inductive analysis was guided by research question 2 (what influenced implementation). The first author (AA) was responsible for coding the material. The codes and initial analyses were discussed in meetings with three of the coauthors (SEO, SD, EL), and this contributed to interpreting, summarizing and synthesizing the data. The inductive analysis resulted in the merging of

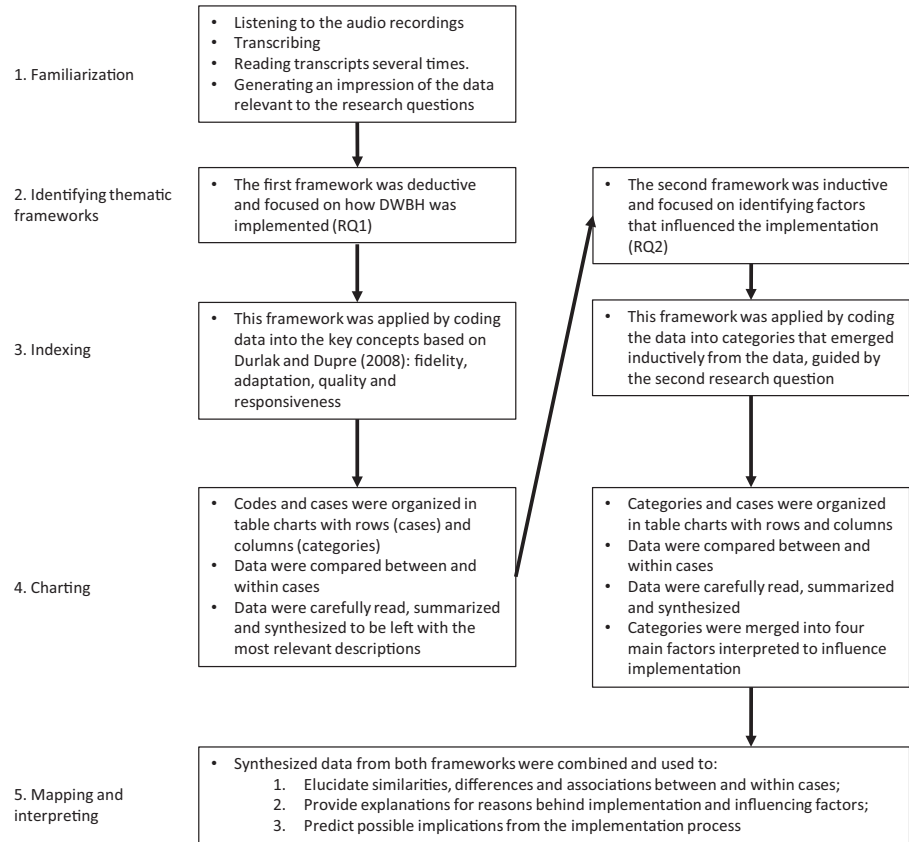


Figure 1.
Detailed outline of how
the five steps of the
framework analysis
were followed

subcategories into four main factors that were interpreted to be the influencers of implementation: (1) frame factors, (2) intervention characteristics, (3) participant characteristics and (4) provider characteristics (see Table 2). The fifth and final step of the analysis (mapping and interpreting) involved combining the inductive and deductive findings to interpret and outline processes behind the implementation and the influencing factors. Notes from the observations assisted in the interpretation and coding of the interview material, in particular by supporting the coherence between what the interviewees stated and what was observed (Mays and Pope, 1995).

Results

In the following section, results are presented within the context of the main factors that were found to influence fidelity, adaptation, quality, responsiveness and dose delivered/received. Average dose delivered in the five schools that were not included in the process evaluation was 80%.

Frame factors

Frame factors represent the contextual opportunities and limitations that exist on an organizational and environmental level. As an influencer to all the examined implementation aspects, scheduling was one of the most influential frame factors. Schools 1 and 4 scheduled DWBH as intended: one lesson was added to the existing schedule, while the other lesson

Main factors	Subcategories	Sample quotes
Frame factors	(1) Adapting intervention components	Subcategory 3, 4 and 5 <i>We do not have access to the gymnasium that one lesson [DW] and we do not have a teacher either, just a substitute teacher sometimes, so we just end up playing cards.</i> (School 3, students)
	(2) After school consequences of DWBH	
	(3) Facilities and equipment	
	(4) Scheduling	
	(5) Teacher availability	
Intervention characteristics	(1) Freedom of choice	Subcategory 1 <i>I think they [students] think that this [DWBH] is very enjoyable, particularly that they are allowed to decide themselves what to do.</i> (School 4, teachers)
	(2) All classes together	
	(3) No assessment	
	(4) Extra time in school	
Participant characteristics	(1) Characteristics, skills, other	Subcategory 1 and 2 <i>How they [students] handle the freedom in DWBH is about maturity. Those who pursue activities in school do what they are supposed to, but some, who probably could be in school, choose something else, off school grounds because it is easier to sneak off.</i> (School 2, teachers)
	(2) Efforts, engagement, interest and motivation	
	(3) Group dynamics, relationships and conflicts	
Provider characteristics	(1) Actions, involvement, supervision, guidance and communication	Subcategory 1 <i>That time when our plans just got out of control, the teacher came and gave us some directions, so we changed the group and changed our plan; he kind of straightened us out.</i> (School 1, students)
	(2) Characteristics, attitudes and motivation	
	(3) Belief in students' capabilities	
	(4) Participating in student activities	

Table 2.
Factors influencing the implementation of "Don't worry, be happy," with merged subcategories and sample quotes

received 5% allocated time from other subjects. School 2 was unable to schedule DWBH as intended and added both lessons on top of the existing schedule, giving students two additional periods in school per week. This had undesirable consequences for the students because they were unable to take the school bus after the extra lessons. Many of these students had long travel distances to school and missing the school bus meant that many arrived at home much later than usual. These consequences strongly influenced how students in School 2 responded to the intervention, illustrated by the following quotes:

School 2, group 1:

Student 1: We do not get a school bus, that's the worst part, I think.

S2: We could have gotten the bus home, because it gets really stressful.

S1: Yes, that's the least they could have done.

(...)

S2: We spend a long time to get home (...) it makes us late for sports practice.

S3: It messes up everything.

School 3 was also unable to schedule the intervention as intended: one subject was removed, and time was allocated from one other subject and one recess period. Although the adaptation limited recess time, the students were still positive toward the intervention.

Schools 1, 2 and 3 scheduled the two lessons on separate days, as the intervention guidelines recommended, while School 4 scheduled DWBH as one double period. The students preferred this adaptation, and according to the teachers, it was necessary:

School 4, teachers:

Teacher 1: Because the gymnasium is a ways away, we would have lost a lot of time if we had two single lessons.

T2: The lessons would have had to be the last of the day (...) so they had time to change clothes before and after. That takes a lot of time.

Facilities emerged as an important factor influencing fidelity, quality, responsiveness and dose received. There were big differences between schools: the gymnasium in School 3 was too small to accommodate two classes at once; therefore, the teacher made the adaptation to have both lessons with separate classes. This scheduling, however, in combination with limited facilities and limited teacher availability, caused another problem for one of the two classes: During DW, they did not have access to the gymnasium because it was being used by another class, and they did not have a teacher because one of the two supervising teachers was on long-term sick leave. Sometimes, a substitute teacher would be present, but they were rarely aware of the purpose of the lesson. The dose delivered in School 3 was registered to be 75%, but these limiting factors suggest that the dose received might be lower:

School 3, teacher:

It was a bit embarrassing last time, because the local ScIM-coordinator came from the university to observe DW, and the substitute had no clue about what was supposed to happen. So the students had just said "well, usually we just play cards in these lessons", so they were just inside the classroom playing cards, which wasn't good.

Schools 1 and 4 had spacious facilities that allowed all the students to participate together. These students expressed satisfaction with the facilities, which included large gymnasiums with ample equipment and many opportunities for outdoor activities. The teachers from

School 2, however, said that their facilities were too limited to carry out DWBH as intended. The gymnasium was small and there was a swimming pool that, according to the students, was often closed for maintenance. The limited facilities led to the adaptation that two-thirds of students had to pursue their activities off school grounds. Additionally, the number of teachers available was limited, as two teachers were always assigned to the swimming pool for safety reasons. These combined factors made it difficult to see and supervise the students on and off school grounds. In contrast, there were always four teachers present during DWBH in School 4. This made it possible to supervise all students inside and outside the gymnasium. The teachers sometimes felt superfluous, but were cognizant of the advantages of being many as illustrated by the following excerpt:

School 4, teachers:

T3: . . .after all, they're 80 students

T2: Yes, and they're spread out, outside and inside

T3: They're everywhere, so if the fire alarm goes off, two teachers will not be enough (. . .) so we might feel superfluous then and there, but it is a safety factor.

T2: Mhm, and when we're four, it's easier to supervise groups well, we can be outside by the soccer field, we can go to the gym down the street. . .

T3: . . .where we have students, and there's a hiking group we can join, so there are more options

Adaptations to the intervention emerged as a factor influencing fidelity, quality and responsiveness. For instance, responsiveness was positively influenced by the adaptation to execute DW in the same way as the BH lesson, which occurred in Schools 1, 3 and 4. In School 2, however, DW was usually organized as what resembled a PE lesson, in which the students could vote on an activity. The teacher justified this adaptation by contending that he had difficult students who did not take DW seriously and needed stricter boundaries. In one class, the vote was often soccer, which left some students unhappy with the lesson:

School 2, group 2:

S1: We have a vote, but the problem is that it always ends up with soccer (. . .) I'm not so fond of soccer; it's okay, but some people hate soccer, so not so many like DW.

Intervention characteristics

Intervention characteristics represent specific components of the intervention, such as additional time in school, additional PA, freedom to choose and the lack of assessment. Various intervention characteristics emerged as factors that influenced responsiveness among students and teachers. Students in Schools 1, 3 and 4 most frequently mentioned the freedom to pursue an activity that interested them as a positive intervention characteristic:

School 1, group 1:

S1: We are interested in it, so you will not get that "oh I do not want to do this" or whatever.

S3: It's like, we do it properly because we like it and then we want to do it.

The students in School 2 did not care that they were able to choose their own activity, because they were not interested in having the intervention at all:

School 2, group 1:

S2: To me, it was just random. I just chose something, to have something to do, really.

(..)

S1: We did not really want to do it (laughter)

Additional time in school also potentially influenced responsiveness. As previously mentioned, students in School 2 received twice as much additional time as DWBH intended, which caused negative responses. Conversely, the students in Schools 1, 3 and 4 received just one additional period. Some of these students took issue with the added time, but also said it was worth it because they could have DWBH.

An important intervention characteristic was that students were not supposed to be assessed during BH. Teachers in Schools 1 and 3, and two student groups, respectively from Schools 3 and 4 talked positively about the lack of assessment. In contrast, a student group from School 2 would have preferred to have the lessons assessed:

School 2, group 2:

S2: It's better if we can, like. . .

S1: Use it for something

S2: Show it, yes, use it for something, like, if you do a great job in BH, none of the teachers see it (..) it's better if it matters for the grade, and it's not just something we do, like, for nothing, since everything is about grades in lower secondary school (laughter).

The students' freedom of choice and the absence of assessment were intervention characteristics that influenced teachers' responsiveness positively in Schools 1, 3 and 4. In contrast, teachers in School 2 acknowledged that DWBH had some good ideas, but argued against some of the intervention's characteristics, such as having BH for all classes simultaneously and allowing students excessive freedom to choose. The following quote illustrates a somewhat negative attitude toward the intervention and the researchers who designed it:

School 2, teachers:

T1: I think the whole thing shows that DWBH was designed by people who do not work in school, because there are a lot of good intentions but when you face common practice, it becomes difficult.

Participant characteristics

Participant characteristics include participants' attitudes, skills, interests, actions and participation that are specifically related to the intervention. Students' attitude toward the intervention was likely an important influencing factor for dose received: all activity groups in Schools 1, 3 and 4 repeatedly stated that they enjoyed DWBH and all groups wanted to continue with the lessons in tenth grade. The delivered doses reported from Schools 1 and 4 were 81 and 86%, respectively, and the positive attitudes expressed in the interviews indicate that the received dose was high, that is, truancy was low. School 2 registered 75% of the dose delivered; however, the interviews indicate that the received dose might have been lower, because of truancy. The negative attitudes toward DWBH suggested high motivation for truancy. Consequently, when BH was scheduled as the last lesson of the day and when two-thirds of students could leave school to do their activities elsewhere, truancy became easy:

School 2, group 2:

S2: There are actually a lot of groups that do not do anything in those lessons (..) I've seen many groups who just go home, or something.

S3: They say they're going on a hike, but they just hike to the bus stop (laughter).

(...)

S1: Yes, I think few actually attend the BH lessons

Interviewer: Have you or anyone else said something about this to the teachers?

S1: No

S2: Snitches get stitches, (laughter)

The participants' interest and motivation also influenced fidelity, in particular regarding the planning process and use of the activity contract. Though participants spent time planning and writing the activity contract, adherence to the activity contract varied within and between schools. In Schools 1, 2 and 4, most groups pursued only one activity and the lessons were more formal than in School 3. In School 3, some groups wrote "various activities" on their activity contract, while others decided on the spot what to do in a given lesson. Some groups also decided to play together, regardless of what they had written in their activity plans. Although the participants initiated this adaptation, an enabling influencing factor was the provider, who allowed it. The adaptation did not seem to have a negative influence on participation or efforts:

School 3, group 2:

S2: Leadership?

S1: No, there's no leadership

S2: We're pretty much, like, we lead each other

S1: We're all leaders, so one of us might say "let's play soccer today", then we play soccer (...) we just run in, get a ball, and start playing.

Conflicts between students emerged as a factor influencing fidelity and was mainly reported in School 4, where they struggled with conflicts within female groups. According to the teachers, DWBH did not cause the conflicts; rather, DWBH helped expose pre-existing hidden conflicts. The conflicts forced the teachers to intervene, because the students could not or would not try to solve the conflicts themselves. This resulted in reduced participation and altered group compositions and a lack of fidelity to groups' conflict resolution plans:

School 4, teachers:

T3: We've definitely steered some students here and there (laughter).

(...)

T2: We're still working on the worst conflicts. But conflicts have been solved by some changing groups.

Provider characteristics

The providers, that is, the teachers, emerged as important influencers of all aspects of implementation. In School 1, one of the two teachers providing DWBH was on long-term sick leave, leaving the second teacher alone with two classes. He expressed a limited ability to supervise all the students because he was mostly compelled to be present in the gymnasium. Sometimes, a substitute teacher was also present, which allowed the main teacher to visit activity groups outside the gymnasium. The teacher maintained that although he could handle being the sole supervisor, two teachers were necessary to ensure all students felt they

were seen and to act as a mediator in case of a conflict between students. Regardless, the remaining teacher's actions and status among the students were positively influencing factors in School 1:

School 1, group 3:

S2: He is the kind of teacher that you have a good relationship with, and you're not afraid to talk to him or anything. He is a really good teacher (. . .) Sometimes he plays soccer with us.

S1: It's fun.

As the quote suggests, the students in School 1 appreciated the teacher's participation in their activity. Similar positive responses occurred in School 3, where the teacher also participated with students on occasion. There were no guidelines from project management regarding teacher participation in the students' activities, but the teacher in School 1 felt that it had a positive influence on his relationship with the students. The students from Schools 1, 3 and 4 spoke of their teachers in either positive or neutral terms. The students from School 2, on the other hand, were more critical of their teachers. For instance, they could not remember what they had written in their activity contract, which they blamed on their teacher. Allegedly, he had collected the contracts when they were completed and the students had not seen them since. One of the groups that had activities off school grounds went even further in its negative description of the teachers:

School 2, group 3:

S2: They definitely do not care about the project. I met a teacher down at the mall after school when I was supposed to be in the Be Happy lesson, and she was like "hi, should not you be in Be Happy?". I was like "yeah", and she just said "okay". Like, the teachers at this school are so bad.

Discussion

The main findings show large differences between schools regarding how DWBH was implemented and how various factors influenced the implementation. Schools 1 and 4 made minor adaptations in the way DWBH was organized, and these were positively received by the students. Intervention characteristics, spacious facilities, scheduling and participant and provider characteristics positively influenced all aspects of implementation. School 2 made major adaptations to how DWBH was scheduled, which reduced both *i* responsiveness and fidelity. Additionally, limited facilities and participant and provider characteristics negatively impacted fidelity, quality and dose received. School 2 was the only school where the intervention was negatively received. School 3 made one major adaptation in how DWBH was organized, and it was poorly received by the students. The intervention itself was otherwise positively received. Limited facilities and scheduling negatively impacted fidelity. Intervention, participant and provider characteristics positively influenced responsiveness, quality and perhaps also dose received.

Reasons for and consequences of adaptations

A common adaptation in Schools 1, 3 and 4 was that the teachers decided to have two identical BH lessons rather than DW and BH. The adaptation was made because the teachers wanted to and because they thought it fit with the purpose of the intervention. Students and teachers agreed that it was a positive adaptation. The adaptations in School 2, however, were mostly made because of contextual limitations. For instance, two periods were added to the schedule instead of one, because all grades and classes followed a fixed schedule, so they could not reorganize the schedule for only some of the students. The adaptation reduced

students' leisure time, which caused a negative response before the intervention had even started. Contextual limitations were the reasons for the main adaptation in School 3 as well and also caused a negative response among the affected students. Initially, limited gymnasium space compelled the teacher to carry out BH in separate classes. However, one of the classes had access to the gymnasium during only one of the two weekly lessons, because it was being used by another PE class. These adaptations and their respective reactions can be elucidated by [Moore et al. \(2013\)](#), who claim that adaptations can be either positive, neutral or negative and either logistical or philosophical. An intervention can be adapted to fit the context and positively influence implementation ([Durlak and Dupre, 2008](#); [Berkel et al., 2011](#)), which is what happened in Schools 1 and 4, and to some extent in School 3, where they made positive adaptations for philosophical reasons. In Schools 2 and 3, however, negative adaptations for logistical reasons negatively impacted implementation. The findings involving adaptations indicate that schools that were likely (because of their preconditions) to succeed in implementing DWBH anyway, made the positive adaptations. Conversely, the schools that were less likely (because of their preconditions) to succeed in implementing DWBH made the negative adaptations.

Dose is not enough

The findings indicated frequent truancy among many students in School 2, caused by negative responsiveness, in combination with poor facilities, scheduling BH to the last lesson of the day and limited teacher supervision. In contrast, positive responsiveness and student interest positively influenced participation in Schools 1, 3 and 4. This concurs with the model designed by [Berkel et al. \(2011\)](#), which proposed that dose received is partly determined by responsiveness. In the review by [Naylor et al. \(2015\)](#), a similar factor to responsiveness, "student characteristics, engagement and motivation," was presented as one of 22 factors influencing dose delivered/received. Although dose received is not an aspect of implementation as defined by [Durlak and Dupre \(2008\)](#), it was included in the review by [Naylor et al. \(2015\)](#), as being one of the most frequently used measures of implementation for school-based PA interventions, along with dose delivered. Dose delivered and dose received provide important information about amount; however, the present study revealed details about the specific components of the intervention that were actually executed, how well they were executed, how students and teachers experienced them and whether these aspects interacted in some way. The present findings suggest that a single focus on dose delivered/received may be a somewhat limited view of implementation that says little about how suitable the intervention is in any school context. This coincides with the findings of [Tarp et al. \(2016\)](#), who conducted a school-based PA intervention cluster RCT and found no effect. According to objective measurements, they could not deliver their target PA dose, but they could not explain why. They therefore recommended that qualitative data on implementation be included in future studies to improve the ability to explain results.

The present results also raise concerns about quantitatively measuring dose delivered without comparing it to dose received. Delivery alone tells us nothing about motivation, interest or actual participation, which has more impact on the intervention outcomes than delivery ([Khanal et al., 2019](#); [Roth, 1985](#); [Durlak and Dupre, 2008](#)). Dose delivered is easy to measure and can be an accurate depiction of the amount provided. However, if dose delivered is the only measure of implementation, researchers may erroneously assume a successful implementation on the grounds of high delivery rates, while low levels of fidelity, quality and responsiveness remain unobserved. On the other hand, and despite plenty of qualitative information, the present results do not tell us how important fidelity, quality and responsiveness to DWBH are for achieving the expected outcomes for physical fitness, mental health, academic achievement or learning environment. The results only indicate that the majority of students and teachers found DWBH to be relevant and enjoyable when (1)

DWBH had few or no consequences for students' leisure time; (2) DWBH was executed with adequate facilities; (3) DWBH was provided by teachers who were present and cared about what the students did; and (4) adaptations were perceived positively and did not negatively impact (1), (2) or (3).

Context and suitability

The present results reveal the complexity of the school context, how schools can differ and how differently schools can carry out a complex and demanding intervention. For instance, facilities were an important factor that differentiated the schools and influenced how DWBH, from the beginning, was adapted, differently, in each school. On top of this, a dynamic interaction between intervention components, teachers and students determined how the implementation process developed. Furthermore, this development occurred differently on multiple levels: the school level (e.g. different facilities, scheduling and teachers), class level (e.g. the scheduling problem in School 3, where one of the participating classes did not have facilities during DW), activity-group level (e.g. few groups within a school pursued the same activities) and student level (e.g. students lost interest and changed groups). [Moore *et al.* \(2019\)](#) argued that introducing a complex intervention in a complex system poses an almost infinite number of uncertainties, which no evaluation is able to address completely. That may be the case, but the process evaluation enables us to address the suitability of DWBH in Norwegian lower secondary schools, which is required in order to say anything about the intervention's generalizability ([Bonell *et al.*, 2006](#)). The somewhat limited implementation that occurred in two out of four schools indicates that DWBH may not have been suitable for these schools. Assuming that other lower secondary schools are as varied as our sample, a strict DWBH program may not be generalizable. However, with the knowledge of the factors that influenced implementation of DWBH, it might be possible to adapt the intervention in a way that fits all contexts, without compromising its purpose. Furthermore, the main negatively influencing factors in Schools 2 and 3 were perhaps not caused by an unsuitable intervention, but by the schools being unable to introduce any program on short notice. Thus, the generalizability of DWBH as a sustainable way to increase PA in lower secondary schools remains uncertain, although we must underscore that having suitable facilities might be the most important precondition.

Design challenges

RCTs are regarded as the gold standard for evaluating effectiveness of public health interventions ([Moore *et al.*, 2015](#); [Victora *et al.*, 2004](#)), and the use of RCTs has been contested ([Byrne, 2013](#)) and defended ([Hawe *et al.*, 2004](#)). Causal inferences from an RCT are based on outcome comparisons between the intervention group and the control group ([Rubin, 1974](#)) and depend on randomization to eliminate differences in observed or unobserved variables between the groups. However, the present results indicate that at least three variables (frame factors, participant characteristics and provider characteristics) varied greatly between schools and may cause systematic differences between intervention group and control group. To avoid these differences, a matched pairs design, based on key frame factors such as facilities, could have been a viable option ([Stuart, 2010](#)). However, in our cluster RCT, the problem was not necessarily differences between intervention group and control group, but large differences within the intervention group. The differences caused DWBH to be implemented differently to the extent that students in different schools received different interventions. DWBH outcomes were most likely influenced by these differences ([Durlak and Dupre, 2008](#)) and randomization, unfortunately, is not a solution to the problem. It is possible that if we had certain predetermined inclusion criteria, such as facilities, we would only have recruited schools that were able to accommodate the intervention. The results might then

have shown fewer implementation differences between the schools. Although this would reduce the representativeness of the included schools, we cannot expect an outcome to change if the intervention school is unable to accommodate the intervention. The variation in facilities may have been the single aspect that mattered the most for the implementation quality, further underscoring the importance of including the aspect in process evaluations. In future cluster RCTs for school-based PA interventions, it is therefore important that researchers ask themselves “what will the intervention require from the school, if it is to be implemented with high quality?” and recruit schools accordingly.

When interpreting the results from a complex cluster-RCT intervention, the results from a process evaluation represent an invaluable tool, and the evaluation should always be conducted whenever there may be variability in the implementation process (Craig *et al.*, 2008; Oakley *et al.*, 2006). Traditionally, the RCT attempts to answer the question “what works?” Combining the RCT with a process evaluation helps us answer “why things work?” (Deaton and Cartwright, 2018) and “under what circumstances?” (Bonell *et al.*, 2012). Answering these questions is essential to designing school-based PA interventions that are feasible for large-scale implementation.

Strengths and limitations

By combining qualitative information on fidelity, adaptation, quality, responsiveness and dose received with quantitative information on dose delivered, the present study provides detailed information regarding the implementation of a school-based PA intervention, compared to previous research (Naylor *et al.*, 2015; Watson *et al.*, 2017; Daly-Smith *et al.*, 2018). Previously conducted process evaluations have highlighted the use of quantitative and qualitative methods as important strengths (De Meij *et al.*, 2013; Burges Watson *et al.*, 2016). However, this study has several limitations. Limited time and resources compelled us to conduct interviews in only four out of nine intervention schools and at only one point in time, rather than multiple times, as recommended by the literature (Moore *et al.*, 2015). As implementations can change over time (Dusenbury *et al.*, 2003), it is important to note that the present findings might not represent implementation throughout the intervention period. Moreover, considering that the qualitative data collection spanned over two months, the implementation process in the first school evaluated might have been at a different stage than in the last school evaluated. Finally, the authors of this paper were stakeholders in the cluster RCT evaluating DWBH and may have an interest in portraying it positively. This may have influenced how we conducted the process evaluation, interpreted and reported the results.

Conclusion

This process evaluation showed that two out of four qualitatively examined schools delivered the intervention with high fidelity, quality, dose delivered and dose received, while obtaining positive responsiveness from participants and providers. The other two schools had major adaptations and limitations and delivered the intervention with varying fidelity, quality, dose delivered and dose received. Frame factors, intervention characteristics, participant characteristics and provider characteristics influenced implementation, and differences between schools may impact the intervention outcomes. Positive adaptations were made in schools that were likely to succeed anyway, based on their preconditions, while negative adaptations were made in schools that, based on their preconditions, were less likely to succeed. The results indicate that adequate facilities and scheduling that did not affect participants' leisure time were important to ensure that the intervention was positively received. Negative responsiveness negatively influenced dose received. Future school-based PA interventions should be designed to generate positive response, perhaps by organizing student-led lessons. However, if responses are negative in certain schools, providers,

researchers and students should cooperate to adapt the intervention in order to make it more relevant and suitable. Careful monitoring of multiple aspects of implementation is key to be able to act upon such responses. We therefore recommend that qualitative process evaluations are conducted on future trials involving school-based PA interventions.

References

- Agans, J.P., Sävfenbom, R., Davis, J.L., Bowers, E.P. and Lerner, R.M. (2013), "Positive movement experiences: approaching the study of athletic participation, exercise, and leisure activity through relational developmental systems theory and the concept of embodiment", *Advances in Child Development and Behavior*, Vol. 45, pp. 261-286.
- Barriball, K.L. and While, A. (1994), "Collecting data using a semi-structured interview: a discussion paper", *Journal of Advanced Nursing-Institutional Subscription*, Vol. 19 No. 2, pp. 328-335.
- Berkel, C., Mauricio, A.M., Schoenfelder, E. and Sandler, I.N. (2011), "Putting the pieces together: an integrated model of program implementation", *Prevention Science*, Vol. 12 No. 1, pp. 23-33.
- Bonell, C., Oakley, A., Hargreaves, J., Strange, V. and Rees, R. (2006), "Assessment of generalisability in trials of health interventions: suggested framework and systematic review", *British Medical Journal*, Vol. 333 No. 7563, pp. 346-349.
- Bonell, C., Fletcher, A., Morton, M., Lorenc, T. and Moore, L. (2012), "Realist randomised controlled trials: a new approach to evaluating complex public health interventions", *Social Science and Medicine*, Vol. 75 No. 12, pp. 2299-2306.
- Burges Watson, D., Adams, J., Azevedo, L.B. and Haighton, C. (2016), "Promoting physical activity with a school-based dance mat exergaming intervention: qualitative findings from a natural experiment", *BMC Public Health*, Vol. 16 No. 1, p. 609.
- Byrne, D. (2013), "Evaluating complex social interventions in a complex world", *Evaluation*, Vol. 19 No. 3, pp. 217-228.
- Clarke, A.M. (2010), "Context matters in programme implementation", *Health Education*, Vol. 110 No. 4, pp. 273-293.
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I. and Petticrew, M. (2008), "Developing and evaluating complex interventions: the new Medical Research Council guidance", *BMJ*, Vol. 337, p. 1655.
- Daly-Smith, A.J., Zwolinsky, S., McKenna, J., Tomporowski, P.D., Defeyter, M.A. and Manley, A. (2018), "Systematic review of acute physically active learning and classroom movement breaks on children's physical activity, cognition, academic performance and classroom behaviour: understanding critical design features", *BMJ Open Sport and Exercise Medicine*, Vol. 4 No. 1, 000341.
- De Meij, J.S.B., Van Der Wal, M.F., Van Mechelen, W. and Chinapaw, M.J.M. (2013), "A mixed methods process evaluation of the implementation of JUMP-in, a multilevel school-based intervention aimed at physical activity promotion", *Health Promotion Practice*, Vol. 14 No. 5, pp. 777-790.
- Deaton, A. and Cartwright, N. (2018), "Understanding and misunderstanding randomized controlled trials", *Social Science and Medicine*, Vol. 210, pp. 2-21.
- Demetriou, Y. and Honer, O. (2012), "Physical activity interventions in the school setting: a systematic review", *Psychology of Sport and Exercise*, Vol. 13 No. 2, pp. 186-196.
- Dobson, D. and Cook, T.J. (1980), "Avoiding type III error in program evaluation: results from a field experiment", *Evaluation and Program Planning*, Vol. 3 No. 4, pp. 269-276.
- Durlak, J.A. (2016), "Programme implementation in social and emotional learning: basic issues and research findings", *Cambridge Journal of Education*, Vol. 46 No. 3, pp. 333-345.
- Durlak, J.A. and Dupre, E.P. (2008), "Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation", *American Journal of Community Psychology*, Vol. 41 Nos 3-4, pp. 327-350.

- Dusenbury, L., Brannigan, R., Falco, M. and Hansen, W.B. (2003), "A review of research on fidelity of implementation: implications for drug abuse prevention in school settings", *Health Education Research*, Vol. 18 No. 2, pp. 237-256.
- Frey, J.H. and Fontana, A. (1991), "The group interview in social research", *The Social Science Journal*, Vol. 28 No. 2, pp. 175-187.
- Gale, N.K., Heath, G., Cameron, E., Rashid, S. and Redwood, S. (2013), "Using the framework method for the analysis of qualitative data in multi-disciplinary health research", *BMC Medical Research Methodology*, Vol. 13 No. 1, p. 117.
- Hawe, P., Shiell, A. and Riley, T. (2004), "Complex interventions: how "out of control" can a randomised controlled trial be?", *BMJ*, Vol. 328 No. 7455, pp. 1561-1563.
- Hawe, P., Shiell, A. and Riley, T. (2009), "Theorising interventions as events in systems", *American Journal of Community Psychology*, Vol. 43 Nos 3-4, pp. 267-276.
- Humphrey, N., Barlow, A. and Lendrum, A. (2017), "Quality matters: implementation moderates student outcomes in the PATHS curriculum", *Prevention Science*, Vol. 19, pp. 197-208.
- Hynynen, S.T., Van Stralen, M.M., Sniehotta, F.F., Araújo-Soares, V., Hardeman, W., Chinapaw, M.J.M., Vasankari, T. and Hankonen, N. (2016), "A systematic review of school-based interventions targeting physical activity and sedentary behaviour among older adolescents", *International Review of Sport and Exercise Psychology*, Vol. 9 No. 1, pp. 22-44.
- Keshavarz, N., 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*, Vol. 70 No. 10, pp. 1467-1474.
- Khanal, S., Choi, L., Innes-Hughes, C. and Rissel, C. (2019), "Dose response relationship between program attendance and children's outcomes in a community based weight management program for children and their families", *BMC Public Health*, Vol. 19 No. 1, p. 716.
- Kolle, E., Solberg, R.B., Grydeland, M., Säfvenbom, R., Anderssen, S.A., Ekelund, U., Berntsen, S., Resaland, G.K., Dyrstad, S.M. and Steene-Johannessen, J. (2020), "Evaluating the effect of increased physical activity in lower secondary school adolescents: the study protocol for the school in motion (ScIM) study – a cluster-randomized controlled trial", Manuscript submitted for publication.
- Kriemler, S., Meyer, U., Martin, E., Van Sluijs, E.M.F., Andersen, L.B. and Martin, B.W. (2011), "Effect of school-based interventions on physical activity and fitness in children and adolescents: a review of reviews and systematic update", *British Journal of Sports Medicine*, Vol. 45 No. 11, pp. 923-930.
- Lerner, R.M. (2015), "Promoting positive human development and social justice: integrating theory, research and application in contemporary developmental science", *International Journal of Psychology*, Vol. 50 No. 3, pp. 165-173.
- Lerner, R.M., Hershberg, R., Hilliard, L. and Johnson, S. (2015), "Concepts and theories in human development", in Bornstein, M.H. and Lamb, M.E. (Eds), *Developmental Science: An Advanced Textbook*, Psychology Press, New York, pp. 3-42.
- Linnan, L. and Steckler, A. (2002), "Process evaluation for public health interventions and research", in Steckler, A. and Linnan, L. (Eds), *Process Evaluation for Public Health Interventions and Research*, Jossey-Bass, San Francisco, pp. 1-23.
- Lubans, D.R., Smith, J.J., Morgan, P.J., Beauchamp, M.R., Miller, A., Lonsdale, C., Parker, P. and Dally, K. (2016), "Mediators of psychological well-being in adolescent boys", *Journal of Adolescent Health*, Vol. 58 No. 2, pp. 230-236.
- Manzano, A. (2016), "The craft of interviewing in realist evaluation", *Evaluation*, Vol. 22 No. 3, pp. 342-360.
- Mays, N. and Pope, C. (1995), "Qualitative Research: observational methods in health care settings", *BMJ*, Vol. 311 No. 6998, pp. 182-184.

- Moore, J., Bumbarger, B. and Cooper, B. (2013), "Examining adaptations of evidence-based programs in natural contexts", *Journal of Primary Prevention*, Vol. 34 No. 3, pp. 147-161.
- Moore, G.F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'cathain, A., Tinati, T., Wight, D. and Baird, J. (2015), "Process evaluation of complex interventions: medical Research Council guidance", *BMJ British Medical Journal*, Vol. 350, p. 7.
- Moore, G.F., Evans, R.E., Hawkins, J., Littlecott, H., Melendez-Torres, G.J., Bonell, C. and Murphy, S. (2019), "From complex social interventions to interventions in complex social systems: future directions and unresolved questions for intervention development and evaluation", *Evaluation*, Vol. 25 No. 1, pp. 23-45.
- Naylor, P.-J., Nettlefold, L., Race, D., Hoy, C., Ashe, M.C., Wharf Higgins, J. and McKay, H.A. (2015), "Implementation of school based physical activity interventions: a systematic review", *Preventive Medicine*, Vol. 72, pp. 95-115.
- Norris, E., Shelton, N., Dunsmuir, S., Duke-Williams, O. and Stamatakis, E. (2015), "Physically active lessons as physical activity and educational interventions: a systematic review of methods and results", *Preventive Medicine*, Vol. 72, pp. 116-125.
- Oakley, A., Strange, V., Bonell, C., Allen, E. and Stephenson, J. (2006), "Process evaluation in randomised controlled trials of complex interventions", *British Medical Journal*, Vol. 332 No. 7538, pp. 413-416.
- Revicki, D.A. and Frank, L. (1999), "Pharmacoeconomic evaluation in the real world", *PharmacoEconomics*, Vol. 15 No. 5, pp. 423-434.
- Rickles, D. (2009), "Causality in complex interventions", *Medicine, Healthcare and Philosophy*, Vol. 12 No. 1, pp. 77-90.
- Robinson, O.C. (2014), "Sampling in interview-based qualitative research: a theoretical and practical guide", *Qualitative Research in Psychology*, Vol. 11 No. 1, pp. 25-41.
- Roth, H.J. (1985), "Relationship between attendance at a parents anonymous adult program and children's behavior at the parents anonymous child care program", *Children and Youth Services Review*, Vol. 7 No. 1, pp. 39-43.
- Rubin, D.B. (1974), "Estimating causal effects of treatments in randomized and nonrandomized studies", *Journal of Educational Psychology*, Vol. 66 No. 5, pp. 688-701.
- Singh, A.S., Saliassi, E., Van Den Berg, V., Uijtendwilligen, L., De Groot, R.H.M., Jolles, J., Andersen, L.B., Bailey, R., Chang, Y.-K., Diamond, A., Ericsson, L., Etner, J.L., Fedewa, A.L., Hillman, C.H., McMorris, T., Pesce, C., Pühse, U., Tomporowski, P.D. and Chinapaw, M.J.M. (2018), "Effects of physical activity interventions on cognitive and academic performance in children and adolescents: a novel combination of a systematic review and recommendations from an expert panel", *British Journal of Sports Medicine*, Vol. 166 No. 1, pp. 49-55.
- Smith, J., Beauchamp, M., Faulkner, G., Morgan, P., Kennedy, S. and Lubans, D.R. (2018), "Intervention effects and mediators of well-being in a school-based physical activity program for adolescents: the 'Resistance Training for Teens' cluster RCT", *Journal of Mental Health and Physical Activity*, Vol. 15, pp. 88-94.
- Spencer, L. and Ritchie, J. (2002), "Qualitative data analysis for applied policy research", in Bryman, A. and Burgess, R.G. (Eds), *Analyzing Qualitative Data*, Routledge, London, pp. 187-208.
- Stuart, E.A. (2010), "Matching methods for causal inference: a review and a look forward", *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, Vol. 25 No. 1, pp. 1-21.
- Tarp, J., Domazet, S.L., Froberg, K., Hillman, C.H., Andersen, L.B. and Bugge, A. (2016), "Effectiveness of a school-based physical activity intervention on cognitive performance in Danish adolescents: LCoMotion—learning, cognition and motion – a cluster randomized controlled trial", *PloS One*, Vol. 11 No. 6, e0158087.
- Victora, C.G., Habicht, J.-P. and Bryce, J. (2004), "Evidence-based public health: moving beyond randomized trials", *American Journal of Public Health*, Vol. 94 No. 3, pp. 400-405.

Watson, A., Timperio, A., Brown, H., Best, K. and Hesketh, K.D. (2017), "Effect of classroom-based physical activity interventions on academic and physical activity outcomes: a systematic review and meta-analysis", *International Journal of Behavioral Nutrition and Physical Activity*, Vol. 14, article 114.

Corresponding author

Andreas Avitsland can be contacted at: andreas.avitsland@uis.no