



REVIEW ARTICLE

For whom and under what circumstances do school-based energy balance behavior interventions work? Systematic review on moderators

MINE YILDIRIM¹, MAARTJE M. VAN STRALEN¹, MAI J. M. CHINAPAW¹, JOHANNES BRUG², WILLEM VAN MECHELEN¹, JOS W. R. TWISK² & SASKIA J. TE VELDE² ON BEHALF OF THE ENERGY-CONSORTIUM

¹Department of Public and Occupational Health and the EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands and ²Department of Epidemiology and Biostatistics and the EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands

Abstract

The aim of this review was to systematically review the results and quality of studies investigating the moderators of school-based interventions aimed at energy balance-related behaviors. We systematically searched the electronic databases of Pubmed, EMBASE, Cochrane, PsycInfo, ERIC and Sportdiscus. In total 61 articles were included. Gender, ethnicity, age, baseline values of outcomes, initial weight status and socioeconomic status were the most frequently studied potential moderators. The moderator with the most convincing evidence was gender. School-based interventions appear to work better for girls than for boys. Due to the inconsistent results, many studies reporting non-significant moderating effects, and the moderate methodological quality of most studies, no further consistent results were found. Consequently, there is lack of insight into what interventions work for whom. Future studies should apply stronger methodology to test moderating effects of important potential target group segmentations.

Key words: Children, diet, intervention, moderator, overweight, physical activity

Introduction

Childhood obesity has risen in both developed and developing countries among both genders and among all ethnic, and socioeconomic groups (1,2). Lower socioeconomic groups and certain ethnic minorities have been hit hardest (3). Obesity is the result of a long-term positive energy balance, i.e., when energy intake is larger than energy expenditure. Intake and expenditure are influenced by so-called energy balance-related behaviors (EBRBs) (4); i.e., specific dietary, physical activity (PA) and sedentary behaviors. The importance of effective interventions that aimed at improving EBRBs in order to prevent obesity in youth was highlighted in previous literature reviews (5,6).

Obesity risk may differ and interventions may not be equally effective across subgroups, such as those

based on socioeconomic status, ethnicity or race, age, and gender (6). One identical intervention strategy (a one-size-fits-all intervention) may not cover the diverse needs of various subgroups; e.g., some may need different types or doses; highly motivated ones may need not more than encouragement (7). Intervention developers can benefit from considering moderating effects by tailoring the intervention content to specific subgroups (8).

Exploring 'for whom' or 'under what circumstances' interventions work or not is possible with moderation analysis. Moderators are variables that affect the direction and/or strength of the relation between independent and outcome variables (9,10). Examples are factors that are manipulated by the intervention (e.g., family involvement), situational (e.g., the site or setting where the intervention is

conducted), socio-demographic (e.g., gender), or psychological variables (e.g., motivation towards behavioural change at the start of the intervention) (10,11).

A moderating effect also called effect modification can be tested by including an interaction term created by multiplying the moderator and the independent variable into the analysis. Conducting subgroup analyses without a previous test of interaction is not advisable, as repeated statistical testing on the same dependent variable for each subgroup increases the risk of obtaining a false positive result (12,13). In case of significant moderation, complementary exploratory analyses within subgroups according to the moderator are needed (14).

To date, only two systematic reviews on moderators of obesity prevention interventions among children have been published. Kremers et al. (15), investigated moderators of intervention effects on EBRBs including only so-called environmental interventions. Stice et al. (16) included studies investigating moderators of obesity intervention effects on overweight indicators. Both reviews found moderating effects of age, gender, and race, but were based on small numbers of studies (15) or lacked studies with long-term follow-up (16). Neither of the reviews performed quality assessments of the included studies or the moderation analyses applied.

The aims of this systematic review were to identify the most important moderators, and to summarize and assess the quality of studies investigating moderators of school-based interventions aimed at EBRBs among school-aged children. We conclude with suggestions for future school-based obesity prevention interventions.

Methods

Literature search

We identified relevant articles through systematic searches in the electronic databases of Pubmed, EMBASE, Cochrane, PsycInfo, ERIC and Sportdiscus. Searches were limited to studies among humans, written in English and published between January 1990 and October 2009. The search terms were based on Boolean logic and included AND-combinations between terms standing for children and adolescents, for school-based intervention and for EBRBs. Since studies are often not framed as a moderation test and generally do not include related terms in their keywords or abstract section, moderator terms were not included in the search strategy. To eliminate an excess number of articles in our broad literature search, NOT-combinations were used for unrelated topics based on our previous experience. The search

strategy for the database of Pubmed is shown in online Figure 1.

Inclusion and exclusion criteria

Inclusion criteria were that the study: (i) had to be a randomized controlled trial (RCT) or quasi-experimental controlled studies aimed at primary prevention of overweight; (ii) targeted EBRBs (PA, sedentary or dietary behaviors) in order to prevent overweight or overweight-related diseases; (iii) was conducted among children and/or adolescents aged between 4 and 18 years; and (iv) applied an appropriate test of moderation (i.e., contained a test of an interaction). Moderators that were included were experimentally manipulated, situational, personal or psychosocial variables. Only full text articles were included. Studies that aimed to change preferences, taste, product sale and content of school lunch were excluded. Studies that were not only school-based (combined with home components) were included in this review.

Process of study selection

Author MY scanned all titles of retrieved studies for relevance. Afterwards, authors MY and MVS independently screened abstracts for possible relevance and decided together on inclusion or exclusion. Next, the same two authors independently checked the full text versions of potentially relevant articles. Authors' differences regarding inclusion were resolved by discussion. A list of excluded studies after full text search and related reasons for exclusion can be obtained from the first author upon request.

Data extraction and quality assessment

All included studies were evaluated and data were abstracted by two authors (MY, MVS) independently and differences resolved by discussion. A third author (MC) was approached in case of disagreements. For each article, the following data were extracted: (i) study population; (ii) study design; (iii) intervention content; (iv) theory that the study was based on; (v) EBRBs outcome variables and measures; (vi) intervention effect on outcome variable; (vii) moderators tested; (viii) results of the moderation analyses; and (ix) in case of significant moderators, results of the stratified analyses.

The quality of the studies was assessed using the items from a Delphi list (17) and from the checklist for evaluating moderation analysis from Frazier et al. (9). Quality items are indicated in online Table II. Among these items, we evaluated point estimates and

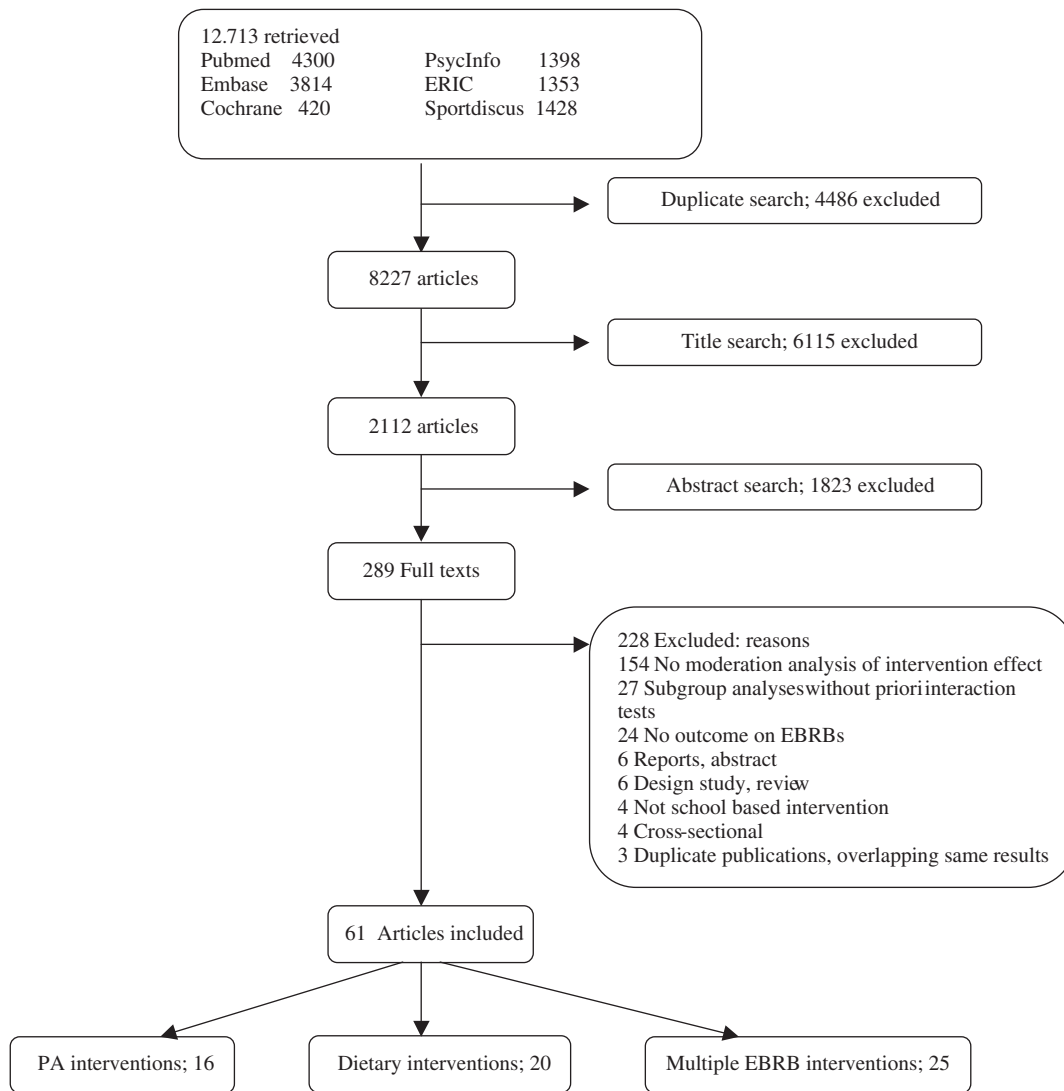


Figure 1. Article search and study selection.

measures of variability as appropriate in case mean, standard deviation, standard error, median or quartiles for the primary outcomes were provided. Criteria have a 'yes' (+ = 1), 'no' (- = 0) or 'don't know' (? = 0) answer format. In case of inadequate information in the text, the reviewers contacted the first author of the study. A quality score was calculated for each study by summing up the scores for each individual quality item, resulting in a possible score of 0–11. The studies were graded arbitrarily as being of a relatively low quality when the total quality score was between 0 and 4, as medium between 5 and 8 and being of a relatively high quality when scores were between 9 and 11.

Moderation analysis requires several conditions that were also evaluated in our quality assessment. First, the selection of moderators should be based on a specific rationale – theory or evidence based – explaining why the intervention may be more effective

for some subgroups than for others. This should be made prior to the intervention planning stage (9). Second, the statistical power of moderation analysis should be considered. Several factors common in social science research can decrease power (e.g., low sample sizes, low effect sizes), which can lead to an incorrect conclusion of no moderating effect (11). Third, homogeneity of residual variances should be checked for categorical moderators. This assumes that the residual variances, which are the error variances that remain after predicting a dependent variable from the independent variables, remain constant across the moderator categories (11,18).

Results

Table I shows the flow of the studies through the review process and the reasons for exclusion. In total 12,713 articles were identified. After scanning titles

and abstracts, we retrieved 289 articles for full text search. Sixty-one articles met our inclusion criteria (19–79). Twenty dietary interventions, 16 PA interventions and 25 multiple EBRB interventions were included. In studies examining only dietary interventions, the majority (16 out of 20) solely aimed to change fruit and vegetable (FV) intake. Multiple EBRB interventions also aimed to improve dietary intake such as fat, sweets and soft drinks consumption and aimed to change PA and sedentary behavior. Some included studies were based on the same intervention (e.g., Squire's Quest, CATCH, Pathways).

Quality assessment

Supplementary Table 2 (available online) shows that 42 studies were of medium and 19 studies were of low quality. There were no high quality studies scoring 9 or higher. The majority of the included studies were RCTs. An intention to treat analysis was applied in 15 studies. Dropout rate was selective in 19 studies, while 26 studies did not report on selectiveness. Twenty-four studies out of the 61 were not explicitly based on a theoretical and/or conceptual framework. Most studies (51 out of 61) used an outcome measure that was checked on its reliability and validity. In most studies, the quality of the conducted moderation analyses was low. Only eight studies provided a rationale for interaction tests, one study checked the assumption of homogeneity of (error) variances across moderator groups. Only one study (43) calculated the power of the moderation analysis and their study had adequate power to detect a moderating effect. We tried to calculate the power using information reported in the articles. However, none of the studies reported sufficient information (i.e. the sample sizes and the predictor-outcome correlations across moderator-based subgroups) for the power calculation program for moderation analysis provided by Aguinis (<http://mypage.iu.edu/~haguinis/mmr/index.html>).

Description of interventions

Supplementary Table 3 (available online) shows the study characteristics. The majority of the studies (19 studies) were conducted in the United States of America (USA), followed by the United Kingdom (six studies), Belgium (five studies), Australia and the Netherlands (both four studies), Norway (three studies), Ireland (two studies) and New Zealand, Cyprus, Greece, Sweden, France, Canada (all one study). One international study was conducted in which the Netherlands, Norway and Spain participated. The number of schools included in the studies ranged from 1–96, and participant numbers ranged

from 122–5106. Three studies were conducted in girls only. Twenty-two studies out of 61 had more than one follow-up while the others had only one follow-up immediately at the end of intervention. Twelve studies used objective measurement methods for PA, such as pedometers and accelerometers. Others used self-reported PA. Lunch observations were used in seven studies, while others used self-report methods for dietary intake. The reliability and validity values of the self-report instruments were mainly lower than the required minimum value of 0.70, as described by Terwee et al. (80).

Intervention effect on outcome

Supplementary Table 3 (available online) reports the intervention effects on EBRBs. The EBRBs evaluated in the studies were total daily PA, light, moderate and vigorous PA, sedentary behavior, active commuting, FV intake, fruit juice consumption, soft drink consumption, dietary fat intake, total energy intake, sugar intake, snacking behavior, sweet consumption and fast food consumption. Studies aimed at increasing FV intake (totally 19) all resulted in significant intervention effects except for one study (24). PA interventions also resulted in significant improvements in children's PA levels especially on moderate to vigorous PA, except two studies (22,28) out of 16. The majority of the multiple EBRB interventions were effective in changing at least one behavior. Five out of 25 studies did not show any significant effects.

Moderators

Supplementary Table 3 (available online) summarizes the results of the moderation analyses conducted in the reviewed studies. The most frequently tested moderating variable was gender (in 46 studies out of 61), followed by ethnicity (in 15 studies), age (in 13 studies), socio-economic status (SES) as indicated by family income and parent's educational level (in 11 studies), baseline value of outcomes (in nine studies) and initial weight status (in eight studies). Other personal moderators were examined each in one study. Psychosocial, situational and experimentally manipulated variables were tested also in a small number of studies.

Personal moderators. The moderating effect of gender was tested in 46 studies (totally 79 interaction tests on multiple behaviors) with a significant moderating effect in 17 studies (21 tests). Moderation analyses showed that in general girls responded significantly better than boys to the interventions, except for the

interventions aimed at changing sedentary behavior and sugar intake (Table I). Girls particularly appeared to respond better to interventions aimed at decreasing their dietary fat intake. Furthermore, baseline levels of outcome behaviors (e.g., baseline FV consumption) moderated the intervention effects on FV consumption, fat intake, snacking and PA in which the groups with unfavorable baseline values benefitted more from the intervention. Age moderated the intervention effect on FV intake in three studies (out of 14). In two studies younger children and in one study older children responded better. The moderation analyses did not yield many statistically significant and consistent results for ethnicity, initial weight status, SES or health status as potential moderators.

Intervention moderators. Family involvement in the intervention was a moderator of the intervention effect on FV consumption with high family involvement showing a larger increase in FV intake (34). Food type provided (fruit vs. vegetables) in the study moderated the intervention effect in two dietary interventions but after stratified analyses, there was no differential intervention effect across subgroups (46,47).

Psychosocial moderators. In one PA intervention children who had lower baseline scores on intention, attitude and perceived behavioral control (PBC) responded better to the intervention (44). Normative beliefs did not moderate the intervention effect in this study.

Situational moderators. Lesson location (57), teacher specialty (57) and compliance (65) to the guided goal setting were moderators of one PA intervention. School food policy (58) and country (72) were moderators of one FV intervention. Type of education (technical vs. normal schools) moderated the intervention effect on fat intake, in which children from technical school responded better to the intervention (38). Average recess time, site (province, state), sports club participation and region did not moderate any intervention effect.

Discussion

The objective of this systematic review was to examine for whom and under what circumstances school-based interventions aimed at EBRBs work. Among the included studies, gender was the most frequently examined variable, followed by ethnicity, age, SES, baseline levels of EBRBs and baseline body weight status. Interventions aimed at changing FV intake,

PA and fat intake were most often affected by the moderators. We did not find a high level of evidence for any of the potential moderators due to the lack of significant and consistent interaction test results. It may, of course, be that intervention effects are not much different for different subgroups, but a possible reason for the lack of moderation found may also be the lack of methodological quality of the included studies. Drop-out analysis, group similarity at baseline and intention-to-treat analysis were not reported or not conducted in more than half of the studies. It was noted by Baron and Kenny (10) that moderation analysis is often applied when there is an unexpected weak or inconsistent relationship between the predictor and the outcome variable. We found that the studies included in the current review were mainly successful (52 studies out of 61) in changing at least one EBRB but still searched for moderating effects. This finding supports our opinion on performing moderation analysis regardless of whether the intervention showed a significant main effect or not. Furthermore, we investigated the possible influence of the methodological quality of the included studies on the obtained moderation analyses results. We found no relationship between study quality and finding significant moderating effects (and also not for finding main effects). Regarding the merits in the design of the studies, some inferences about potential moderators can be derived.

The most convincing evidence was found for the moderating effect of gender; mainly girls responded significantly better to the interventions than boys. The review from Kremers et al. (15) focused on environmental interventions to change EBRBs and found similar results regarding moderating effects of gender, mainly due to girls' better response to the interventions. One possible explanation for this could be that boys are generally more physically active than girls leaving more room for improvement among girls (81). This finding has been linked to differences in motor skills development, body composition, socialization towards sports and physical activity and freedom to involve to activities independently outside the home (82). One way to explore the factors that prevent the success of a school-based intervention in the full range of student is the assessment of process measures. Although a process evaluation was conducted in nine out of 31 studies in which moderating effects were found, only one study (38) conducted a process evaluation across subgroups. The authors found that girls had read the intervention messages significantly more often than boys and girls found the instructions significantly clearer than boys. These results suggest that the intervention was better implemented among girls. There was no information on differences in validity and reliability of

outcome measures between girls and boys in the included studies. Nevertheless, Rangul et al. (83) showed that the reliability of a frequently used self-administered questionnaire (WHO-HBSC) to measure PA was significantly higher in girls compared to boys and that facilitates finding an intervention effect in girls rather than in boys. It may also be the case that girls are more likely to respond in a socially desirable way to self-report questionnaires, such as on activity or fruit and vegetable intake. In future interventions, gender specific underlying mechanisms of behavior change should therefore be considered. For instance, Simen-Kapeu et al. (84) found that girls have fewer active role models, more barriers and less perceived benefits of physical activity compared to boys. Future interventions may consider targeting their intervention strategies more to gender characteristics. For instance, favorable effects of active gaming and organized competitive, team, and high intensity sports and exercises (e.g., football, basketball, team handball, cycling) on PA that are popular among boys, have been shown (85,86). Regarding the dietary interventions, girls have higher concerns of body weight gain and body image when compared to boys (87). This can result in a higher interest towards the intervention increasing the likelihood of an intervention effect. In a qualitative study, it was shown that the key motivating factors for boys in terms of healthy eating were sports and physical performance (88).

The second most common moderator was the baseline level of EBRBs (e.g., baseline FV consumption). These moderators particularly showed their moderating effect in dietary interventions. The subgroups with unfavorable baseline values responded better to the intervention on FV consumption, fat intake and snacking. The results indicated that the more the children needed the intervention, or the more room there was for improvement, the more they appeared to benefit from it. Regression to the mean – the fact that participants with extreme scores at baseline generally regress to the population mean of a group – should be considered and avoided in future studies.

In several papers, the importance of parental involvement in childhood overweight prevention interventions has been argued (89,90). Its moderating effect was tested in one multiple behavior intervention (34) and found as a significant moderator of the intervention effect on FV intake but not on total PA and sedentary behavior. In this particular study on FV intake, children with parents who were highly involved in the intervention program responded better to the intervention. This is in line with a previous review on parental involvement. Hingle et al. (91) showed that directly involving parents to dietary

interventions showed promising results. Despite this finding, the current review provides insufficient evidence on the moderating effects of parental involvement to draw strong conclusions.

In contrast to general expectations, baseline body weight status (e.g., obese vs. normal), ethnicity, SES and age did not yield remarkable, consistent moderating effects. Considering the fact that the lower SES and ethnic minority children have higher risk for obesity and undesired EBRB habits, more high quality research is especially needed to look at the moderating effect of SES and ethnicity. Undoubtedly, the preferred public health option is not having a moderating effect, but a good quality intervention that is highly effective across different subgroups. Due to the lack of information concerning the power of the applied moderation analyses we cannot state with confidence that lack of significant moderators in our review indeed means lack of existing moderating effects.

The psychosocial variables (i.e., intention, attitude and perceived behavior control) moderated the intervention effect on PA level in one study in which children benefitted more who had lower baseline values on these variables. Since these moderators were only analyzed in one study, the generalizability of these results is limited.

Methodological issues

The overall quality of the conducted moderation analysis of the studies included in this review was unsatisfactory and the methodology needs more careful considerations in future studies. First, many studies conducted stratified analysis of the intervention effect within each subgroup without an appropriate interaction test ($n=27$). This major shortcoming was a reason to exclude studies from our review. Testing several subgroups simultaneously increases the probability of finding significant results due to chance alone (13,92). For example, categorizing age into four age groups yields four statistical tests to examine the intervention effect on the same outcome variable (92), increasing the likelihood of finding a Type I error (false positive result). The probability of one or more false positive results is about 5% in one test, and increases to 10% for two tests (e.g., two gender groups), and to 14% for three tests (e.g., three age groups) (92). In contrast, using a single interaction term produces a single test and should therefore be favored above subgroup analyses (92).

Second, the fact that many studies did not find a moderating effect, could be explained by a lack of important moderators or a too low power in the studies in order to be able to detect a moderating effect.

Moderator analyses often have poor statistical power. Factors that determine the power of a study to detect interactions are the sample size, the size of the moderating effect, the equality of subgroup sizes, measurement error in the variables that constitute the interaction term and categorization of a truly continuous variable (11,93). Categorization of continuous variables leads also to loss of effect sizes and statistical power, and it would likely result in the loss of information about individual differences within the groups (94). Conducting a power analysis before the start of the study helps researchers to maximize statistical power by design and measurement choices. Aguinis (11) provided the statistical software programs for power calculation for categorical moderators. Jaccard et al. (14) provided tables for estimating power for interactions. The most certain strategy to increase power is recruiting the largest sample available (95). This can cause a very large sample size and can be impractical. Another alternative is choosing measures with high reliability since measurement error of variables negatively influences the estimated effect sizes for the interaction term and consequently the power (9,11). Another strategy is to increase Type I error (α) rates by increasing their p -value from the typical 0.05 to for example 0.10 when conducting a moderation analysis (95). This is also preferable due to the fact that moderating effects are often small. Looking at p -values, however, is risky, since accepting the null hypothesis when statistical significance is not found has created difficulty in bringing out interaction effects (96).

The third methodological issue that needs improvement is related to the homogeneity of error variance that has an effect on p -values of the interaction tests. Homogeneity of error variance is reflected in the distribution of the residuals of the regression analyses. Among the included studies only one study (66) checked the homogeneity of error variance. Researchers should be aware of the fact that when they violate the assumption of homogeneity, the chance of finding a Type I (false positive result) or Type II (false negative result) error increases, depending on the specific sample. Consequently, the results of moderation analyses cannot be trusted (11,18). Homogeneity can be checked with a visual examination of the distributions of residuals along the regression line in a simple scatter plot separately for moderator groups, by a statistical web-based program provided by Aguinis (11) or by a Levene test in ANOVA. When the error variances are not homogenous, the weighted least square approach can be applied as an alternative to ordinary least square regression. This approach corrects the heterogeneity of error variances by obtaining a single weight for each group (97). In addition, applying structural

equation modelling (SEM) is a good alternative since SEM analyses do not require the assumption of homogeneity (11,95).

Fourth, potential moderators should be based on a theory, prior findings or literature reviews. In the current literature review we found that only a very small number of studies fulfilled this requirement. When researchers conduct moderation analyses without a rationale, they may be tempted to analyze all variables available as potential moderators, this will increase the chance of finding false positive results (98).

Limitations and strengths

Our findings should be interpreted in light of methodological strengths and weaknesses. The strength of our review was the extensiveness of literature search as well as being the first paper that assessed the quality of moderation analysis in school based-interventions aimed at EBRBs among youth. We also evaluated the methodological quality of included studies, which was not done by the previous review by Kremer et al. (15). Although our review covered a wide range of school-based obesity prevention interventions, most studies were limited mainly in terms of their methodological quality.

Recommendations for future studies

Methodological implications

By including a large enough sample size, determined *a priori* by a power calculation, including balanced subgroups, and using reliable measures with continuous scales (rather than artificially dichotomized scales) researchers may increase the chance of correctly estimating a moderating effect. Reporting moderation analysis in the abstract would increase knowledge building and enable easy access to the study by other researchers aimed at exploring moderators (98).

Researchers who want to explore moderation effects should restrict their investigations to a specific rationale and avoid analyzing variables without a rationale. One other important issue is reporting/publication bias. Researchers may have applied many tests but report only the significant ones. Another line of research is applying SEM in order to explore moderation effects in studies including latent variables measured by multiple indicators. This statistical method has, when compared to the more traditional regression analyses, the advantage on power that it corrects for measurement error (95,99). An example is reported by Hopwood (100).

Theoretical implications

The potential moderators stimulate a search for theories and underlying mechanisms why an intervention has differential effects. Moderating effects could also serve to test an underlying theoretical model. Current behavioral change theories do not consider moderating effects that need to be taken into account in the development or improvement of theories. Regarding the importance of contribution of psychosocial factors in health behaviors, information is necessary concerning moderation effects of psychosocial variables. A greater understanding of why different subgroups respond in different ways to an intervention should be further explored. Different subgroups may benefit from different intervention strategies. Therefore, future investigations for varying types of intervention strategies need to be conducted in both qualitative and quantitative studies. Potential moderation effects, such as gender differences in intervention effects, suggest the need for separate programs or at least special considerations for boys and girls that can be provided by tailored intervention messages.

Conclusion

This systematic review leads to two conclusions. First, the systematic review cannot be viewed as conclusive due to the inconsistent results found in the included studies and the small numbers of significant moderation effects. However, gender (female) and unfavorable baseline values are the most prominent moderators of the intervention effects. The further investigation of underlying mechanisms will help inform the delivery of interventions to those who might benefit the most. Second, the overall methodological quality of the included studies was moderate. The quality of the moderation analysis needs improvement in future studies.

Acknowledgements

The contribution of MY was funded by the World Cancer Research Fund (WCRF grant 2008/65), the contribution of MVS was funded by the Netherlands Organization for Health Research and Development (ZonMw 121.520.002). The ENERGY-project is funded by the Seventh Framework Programme (CORDIS FP7) of the European Commission, HEALTH (FP7-HEALTH-2007-B). The content of this article reflects only the authors' views and the European Community is not liable for any use that may be made of the information contained therein.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Flynn MTA, McNeil DA, Maloff B et al. Reducing obesity and related chronic disease risk in children and youth: a synthesis of evidence with 'best practice' recommendations. *Obes Rev.* 2006;7(Suppl. 1):7–66.
2. Cali AMG, Caprio S. Obesity in children and adolescents. *J Clin Endocrinol Metab.* 2008;93(11):31–6.
3. Wang Y, Beydoun MA. The obesity epidemic in the United States—gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis. *Epidemiol Rev.* 2007;29:6–28.
4. Kremers SPJ, Visscher TLS, Seidell JC, van Mechelen W, Brug J. Cognitive determinants of energy balance-related behaviours: measurement issues. *Sports Med.* 2005;35(11):923–33.
5. Brown T, Summerbell C. Systematic review of school-based interventions that focus on changing dietary intake and physical activity levels to prevent childhood obesity: an update to the obesity guidance produced by the National Institute for Health and Clinical Excellence. *Obes Rev.* 2009;10:110–41.
6. Maziak W, Ward KD, Stockton MB. Childhood obesity: are we missing the big picture? *Obes Rev.* 2008;9:35–42.
7. Baranowski T, Cerin E, Baranowski J. Steps in the design, development and formative evaluation of obesity prevention-related behavior change trials. *Int J Behav Nutr Phys Act.* 2009;6:6.
8. Donaldson SI. Mediator and moderator analysis in program development. In Sussman S, editor. *Handbook of Program Development for Health Behaviour Research and Practice.* Thousand Oaks, CA: Sage Publications, Inc. p. 470–500.
9. Frazier PA, Barron KE, Tix AP. Testing moderator and mediator effects in counseling psychology. *J Couns Psychol.* 2004;51(1):115–34.
10. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol.* 1986;51(6):1173–82.
11. Aguinis H. *Regression analysis for categorical moderators.* 1st ed. New York: The Guilford Press; 2004.
12. Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol.* 2004;57:229–36.
13. Cui L, Hung HMJ, Wang SJ, Tsong Y. Issues related to subgroup analysis in clinical trials. *J Biopharmaceut Stat.* 2002;12(3):347–58.
14. Jaccard J, Turrissi R, Wan CK. *Interaction effects in multiple regression.* 1st ed. Newbury Park, CA: Sage; 1990.
15. Kremers SPJ, de Bruijn GJ, Droomers M, van Lenthe F. Moderators of environmental intervention effects on diet and activity in youth. *Am J Prev Med.* 2007;32(2):163–72.
16. Stice E, Shaw H, Marti N. A meta-analytic review of obesity prevention programs for children and adolescents: the skinny on interventions that work. *Psychol Bull.* 2006;132(5):667–91.
17. Verhagen AP, de Vet HC, de Bie RA et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol.* 1998;51(12):1235–41.

18. Fairchild AJ, McQuillin SD. Evaluating mediation and moderation effects in school psychology: a presentation of methods and review of current practice. *J School Psychol.* 2010;48:53–84.
19. Ashfield-Watt PAL, Stewart EA, Scheffer JA. A pilot study of the effect of providing daily free fruit to primary-school children in Auckland, New Zealand. *Pub Health Nutr.* 2008;12(5):693–701.
20. Baranowski T, Davis M, Resnicow K et al. Gimme 5 fruit, juice, and vegetables for fun and health: outcome evaluation. *Health Educ Behav.* 2000;27(1):96–111.
21. Baranowski T, Baranowski J, Cullen KW et al. Squire's quest! Dietary outcome evaluation of a multimedia game. *Am J Prev Med.* 2003;24(1):52–61.
22. Barnett LM, van Beurden E, Morgan PJ, Brooks LO, Zask A, Beard JR. Six year follow-up of students who participated in a school-based physical activity intervention: a longitudinal cohort study. *Int J Behav Nutr Phys Act.* 2009;6.
23. Bere E, Veierod MB, Bjelland M, Klepp KI. Free school fruit – sustained effect 1 year later. *Health Educ Res.* 2006; 21(2):268–75.
24. Bere E, Veierod MB, Bjelland M, Klepp KI. Outcome and process evaluation of a Norwegian school-randomized fruit and vegetable intervention: fruits and vegetables make the marks (FVMM). *Health Educ Res.* 2006;21(2):258–67.
25. Bere E, Veierod MB, Klepp KI. The Norwegian School Fruit Programme: evaluating paid vs. no-cost subscriptions. *Prev Med.* 2005;41(2):463–70.
26. Burke V, Thompson C, Taggart AC et al. Differences in response to nutrition and fitness education programmes in relation to baseline levels of cardiovascular risk in 10 to 12-year-old children. *J Hum Hypertens.* 1996;10(Suppl. 3): 99–106.
27. Butcher Z, Fairclough S, Stratton G, Richardson D. The effect of feedback and information on children's pedometer step counts at school. *Pediatr Exerc Sci.* 2007; 19(1):29–38.
28. Cardon G, Labarque V, Smits D, Bourdeaudhuij ID. Promoting physical activity at the preschool playground: the effects of providing markings and play equipment. *Prev Med.* 2009;48(4):335–40.
29. Cullen KW, Watson K, Baranowski T, Baranowski JH, Zakeri I. Squire's quest: intervention changes occurred at lunch and snack meals. *Appetite.* 2005;45(2):148–51.
30. Dziewaltowski DA, Estabrooks PA, Welk G et al. Healthy youth places: a randomized controlled trial to determine the effectiveness of facilitating adult and youth leaders to promote physical activity and fruit and vegetable consumption in middle schools. *Health Educ Behav.* 2009;36(3): 583–600.
31. Ernst MP, Pangrazi RP. Effects of a physical activity program on children's activity levels and attraction to physical activity. *Pediatr Exerc Sci.* 1999;11(4):393–405.
32. Foster GD, Sherman S, Borradaile KE et al. A policy-based school intervention to prevent overweight and obesity. *Pediatrics.* 2008;121(4):794–802.
33. Frenn M, Malin S, Bansal N et al. Addressing health disparities in middle school students' nutrition and exercise. *J Community Health Nurs.* 2003;20(1):1–14.
34. Gentile DA, Welk G, Eisenmann JC et al. Evaluation of a multiple ecological level child obesity prevention program: switch what you do, view, and chew. *BMC Med.* 2009;7(49): 1–12.
35. Going S, Thompson J, Cano S et al. The effects of the Pathways Obesity Prevention Program on physical activity in American Indian children. *Prev Med.* 2003;37(6 Pt 2):62–9.
36. Goran MI, Reynolds K. Interactive multimedia for promoting physical activity (IMPACT) in children. *Obes Res.* 2005;13(4):762–71.
37. Graham DJ, Schneider M, Cooper DM. Television viewing: moderator or mediator of an adolescent physical activity intervention? *Am J Health Promot.* 2008 Nov;23(2):88–91.
38. Haerens L, Deforche B, Maes L, Brug J, Vandelanotte C, De Bourdeaudhuij I. A computer tailored dietary fat intake intervention for adolescents: results of a randomized controlled trial. *Ann Behav Med.* 2007;34(3):253–62.
39. Haerens L, Deforche B, Vandelanotte C, Maes L, De Bourdeaudhuij I. Acceptability, feasibility and effectiveness of a computer-tailored physical activity intervention in adolescents. *Patient Educ Couns.* 2007;66(3):303–10.
40. Haerens L, De Bourdeaudhuij I, Maes L, Vereecken C, Brug J, Deforche B. The effects of a middle-school healthy eating intervention on adolescents' fat and fruit intake and soft drinks consumption. *Pub Health Nutr.* 2007;10(5): 443–9.
41. Haerens L, De Bourdeaudhuij I, Maes L, Cardon G, Deforche B. School-based randomized controlled trial of a physical activity intervention among adolescents. *J Adolesc Health.* 2007;40(3):258–65.
42. Harrison M, Murphy N, Burns C, Heslin J, McGuinness M. Influence of a health education intervention on physical activity and screen inactivity in Irish primary school children: switch off-get active. (Abstract). *J Sci Med Sport.* 2005; 8(4 Suppl.):41.
43. Hendy HM, Williams KE, Camise TS, Alderman S, Ivy J, Reed J. Overweight and average weight children equally responsive to 'Kids Choice Program' to increase fruit and vegetable consumption. *Appetite.* 2007;49(3):683–6.
44. Hill C, Abraham C, Wright DB. Can theory-based messages in combination with cognitive prompts promote exercise in classroom settings? *Soc Sci Med.* 2007;65(5):1049–58.
45. Himes JH, Ring K, Gittelsohn J et al. Impact of the Pathways intervention on dietary intakes of American Indian school-children. *Prev Med.* 2003;37(6 Pt 2):55–61.
46. Horne PJ, Tapper K, Lowe CF, Hardman CA, Jackson MC, Woolner J. Increasing children's fruit and vegetable consumption: a peer-modelling and rewards-based intervention. *Eur J Clin Nutr.* 2004;58(12):1649–60.
47. Horne PJ, Hardman CA, Lowe CF et al. Increasing parental provision and children's consumption of lunchbox fruit and vegetables in Ireland: the Food Dudes intervention. *Eur J Clin Nutr.* 2009;63(5):613–8.
48. Kelder SH, Perry CL, Peters RJ, Lytle LL, Klepp KI. Gender differences in the class of 1989 study: the school component of the Minnesota Heart Health Program. *J Health Educ.* 1995;26(2):36–44.
49. Kipping RR, Payne C, Lawlor DA. Randomised controlled trial adapting US school obesity prevention to England. *Arch Dis Child.* 2008;93(6):469–73.
50. Loucaides CA, Jago R, Charalambous I. Promoting physical activity during school break times: piloting a simple, low cost intervention. *Prev Med.* 2009;48(4):332–4.
51. Lowe CF, Horne PJ, Tapper K, Bowdery M, Egerton C. Effects of a peer modelling and rewards-based intervention to increase fruit and vegetable consumption in children. *Eur J Clin Nutr.* 2004;58(3):510–22.
52. Luepker RV, Perry CL, McKinlay SM et al. Outcomes of a field trial to improve children's dietary patterns and physical activity. The Child and Adolescent Trial for Cardiovascular Health (CATCH) collaborative group. *JAMA.* 1996;275(10): 768–76.
53. Manios Y, Kafatos A. Health and nutrition education in elementary schools: changes in health knowledge, nutrient

- intakes and physical activity over a six year period. *Pub Health Nutr.* 1999;2(3A):445–8.
54. Manios Y, Moschandreas J, Hatzis C, Kafatos A. Evaluation of a health and nutrition education program in primary school children of Crete over a three-year period. *Prev Med.* 1999;28(2):149–59.
 55. Marcus C, Nyberg G, Nordenfelt A, Karpmyr M, Kowalski J, Ekelund U. A 4-year, cluster randomized, controlled childhood obesity prevention study: STOPP. *Int J Obes.* 2009; 33(4):408–17.
 56. Martens MK, van Assema P, Paulussen TG, van Breukelen G, Brug J. Krachtvoer: effect evaluation of a Dutch healthful diet promotion curriculum for lower vocational schools. *Pub Health Nutr.* 2008;11(3):271–8.
 57. McKenzie TL, Stone EJ, Feldman HA et al. Effects of the CATCH physical education intervention: teacher type and lesson location. *Am J Prev Med.* 2001;21(2):101–9.
 58. Moore L, Tapper K. The impact of school fruit tuck shops and school food policies on children's fruit consumption: a cluster randomised trial of schools in deprived areas. *J Epidemiol Comm Health.* 2008;62(10):926–31.
 59. Perry CL, Bishop DB, Taylor G et al. Changing fruit and vegetable consumption among children: the 5-a-Day Power Plus program in St. Paul, Minnesota. *Am J Pub Health.* 1998;88(4):603–9.
 60. Perry CL, Lytle LA. Effects of the child and adolescent trial for cardiovascular health (CATCH) on fruit and vegetable intake. *J Nutr Educ.* 1998;30(6):354–60.
 61. Perry CL, Bishop DB, Taylor GL et al. A randomized school trial of environmental strategies to encourage fruit and vegetable consumption among children. *Health Educ Behav.* 2004;31(1):65–76.
 62. Reinaerts E, de Nooijer J, Candel M, de Vries N. Increasing children's fruit and vegetable consumption: distribution or a multicomponent programme? *Pub Health Nutr.* 2007; 10(9):939–47.
 63. Robinson TN. Reducing children's television viewing to prevent obesity: a randomized controlled trial. *JAMA.* 1999;282(16):1561–7.
 64. Salmon J, Ball K, Hume C, Booth M, Crawford D. Outcomes of a group-randomized trial to prevent excess weight gain, reduce screen behaviours and promote physical activity in 10-year-old children: switch-play. *Int J Obes (Lond).* 2008;32(4):601–12.
 65. Shilts MK, Horowitz M, Townsend MS. Guided goal setting: effectiveness in a dietary and physical activity intervention with low-income adolescents. *Int J Adolesc Med Health.* 2009;21(1):111–22.
 66. Simon C, Schweitzer B, Oujaa M et al. Successful overweight prevention in adolescents by increasing physical activity: a 4-year randomized controlled intervention. *Int J Obes (Lond).* 2008;32(10):1489–98.
 67. Singh AS, Chin AP, Brug J, Van Mechelen W. Dutch obesity intervention in teenagers: effectiveness of a school-based program on body composition and behavior. *Arch Pediatr Adolesc Med.* 2009;163(4):309–17.
 68. Stevens J, Story M, Ring K, Murray DM, Cornell CE, Gittelsohn J. The impact of the Pathways intervention on psychosocial variables related to diet and physical activity in American Indian schoolchildren. *Prev Med.* 2003;37 (6 Pt 2):70–9.
 69. Stock S, Miranda C, Evans S et al. Healthy buddies: a novel, peer-led health promotion program for the prevention of obesity and eating disorders in children in elementary school. *Pediatrics.* 2007;120(4):1059–68.
 70. Tak NI, te Velde SJ, Brug J. Ethnic differences in 1-year follow-up effect of the Dutch Schoolgruitem Project – promoting fruit and vegetable consumption among primary-school children. *Pub Health Nutr.* 2007;10(12):1497–507.
 71. Tak NI, te Velde SJ, Brug J. Long-term effects of the Dutch Schoolgruitem Project – promoting fruit and vegetable consumption among primary-school children. *Pub Health Nutr.* 2009;12(8):1213–23.
 72. Te Velde SJ, Brug J, Wind M et al. Effects of a comprehensive fruit- and vegetable-promoting school-based intervention in three European countries: the Pro Children Study. *Br J Nutr.* 2008;99(4):893–903.
 73. Vandongen R, Jenner DA, Thompson C et al. A controlled evaluation of a fitness and nutrition intervention program on cardiovascular health in 10- to 12-year-old children. *Prev Med.* 1995;24(1):9–22.
 74. Verstraete S, Cardon G, De Clercq D, De Bourdeaudhuij I. Effectiveness of a two-year health-related physical education intervention in elementary schools. *J Teach Phys Educ.* 2007;26(1):20–34.
 75. Verstraete SJM, Cardon GM, de Clercq DLR, de Bourdeaudhuij IMM. Increasing children's physical activity levels during recess periods in elementary schools: the effects of providing game equipment. *Eur J Pub Health.* 2006;16(4): 415–9.
 76. Verstraete SJM, Cardon GM, de Clercq DLR, de Bourdeaudhuij IMM. A comprehensive physical activity promotion programme at elementary school: the effects on physical activity, physical fitness and psychosocial correlates of physical activity. *Pub Health Nutr.* 2007;10(5):477–84.
 77. Webber LS, Catellier DJ, Lytle LA et al. Promoting physical activity in middle school girls: trial of activity for adolescent girls. *Am J Prev Med.* 2008;34(3):173–84.
 78. Wechsler H, Basch CE, Zyberty P, Shea S. Promoting the selection of low-fat milk in elementary school cafeterias in an inner-city Latino community: evaluation of an intervention. *Am J Pub Health.* 1998;88(3):427–33.
 79. Young DR, Phillips JA, Yu T, Haythornthwaite JA. Effects of a life skills intervention for increasing physical activity in adolescent girls. *Arch Pediatr Adolesc Med.* 2006;160(12): 1255–61.
 80. Terwee CB, Bot SDM, de Boer MR et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34–42.
 81. Biddle SJH, Gorely T, Stensel DJ. Health-enhancing physical activity and sedentary behaviour in children and adolescents. *J Sport Sci.* 2004;22:679–701.
 82. Sweeting HN. Gendered dimensions of obesity in childhood and adolescence. *Nutr J.* 2008;7:1–14.
 83. Rangul V, Holmen TL, Kurtze N, Cuypers K, Midthjell K. Reliability and validity of two frequently used self-administered physical activity questionnaires in adolescents. *BMC Med Res Meth.* 2008;8:47.
 84. Simen-Kapeu A, Veugelers PJ. Should public health interventions aimed at reducing childhood overweight and obesity be gender-focused? *BMC Pub Health.* 2010;10:340.
 85. Graves L, Stratton G, Ridgers ND, Cable NT. Energy expenditure in adolescents playing new generation computer games. *BMJ.* 2007;335:1282–4.
 86. Vilhjalmsdottir R, Kristjansdottir G. Gender differences in physical activity in older children and adolescents: the central role of organized sport. *Soc Sci Med.* 2003;56(2):363–74.
 87. Davison KK, Markey CN, Birch LL. A Longitudinal examination of patterns in girls' weight concerns and body dissatisfaction from ages 5 to 9 years. *Int J Eat Disord.* 2003; 33(3):320–32.
 88. McKinley MC, Lewis C, Robson PJ et al. It's good to talk: children's views on food and nutrition. *Eur J Clin Nutr.* 2005;59:542–51.

89. Kremers SPJ, de Bruijn GJ, Visscher TLS, van Mechelen W, de Vries NK, Brug H. Environmental influences on energy balance-related behaviours: a dual-process view. *Int J Bev Nutr Phys Act.* 2006;3:9.
90. Bautista-Castan I, Doreste J, Serra-Majem L. Effectiveness of interventions in the prevention of childhood obesity. *Eur J Epid.* 2004;19:617–22.
91. Hingle MD, O'Connor TM, Dave JM, Baranowski T. Parental involvement in interventions to improve child dietary intake: a systematic review. *Prev Med.* 2010;51:103–111.
92. Schulz KF, Grimes DA. Multiplicity in randomised trials II: subgroup and interim analyses. *Lancet.* 2005;365:1657–61.
93. Villa JR, Howell JP, Dorfman PW, Daniel DL. Problems with detecting moderators in leadership research using moderated multiple regression. *The Leadership Q.* 2003;14:3–23.
94. MacCallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. *Psychol Meth.* 2002;7(1):19–40.
95. Stone-Romero EF, Liakhovitski D. Strategies for detecting moderator variables: a review of conceptual and empirical issues. *Res Personnel Human Resources Manage.* 2002;21:333–72.
96. Cronbach LJ. Statistical tests for moderator variables: flaws in analyses recently proposed. *Psychol Bull.* 1987;102(3):414–7.
97. Overton RC. Moderated multiple regression for interactions involving categorical variables: A statistical control for heterogeneous variance across two groups. *Psychol Meth.* 2001;6(3):218–33.
98. Petrosino A. Mediators and moderators in the evaluation of programs for children: current practice and agenda for improvement. *Eval Rev.* 2000;24:47–71.
99. McClelland GH, Judd CM. Statistical difficulties of detecting interactions and moderator effects. *Psychol Bull.* 1993;114(2):376–90.
100. Hopwood CJ. Moderation and mediation in structural equation modeling: applications for early intervention research. *J Early Intervent.* 2007;29(3):262–72.

Supplementary material available online

Supplementary Tables 1–3.