PART 2

Digging deep, thinking hard on effects of school-based interventions targeting energy balance-related behaviors

3

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

Mine Yıldırım, Maartje M van Stralen, Mai JM Chinapaw,
Johannes Brug, Willem van Mechelen,
Jos WR Twisk and Saskia J te Velde

International Journal of Pediatric Obesity,
2011;6(2-2):e46-57

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.

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 behavioral 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 supplementary Table 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 supplementary Table 2. Among these items, we evaluated point estimates and 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

Figure 1 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 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).

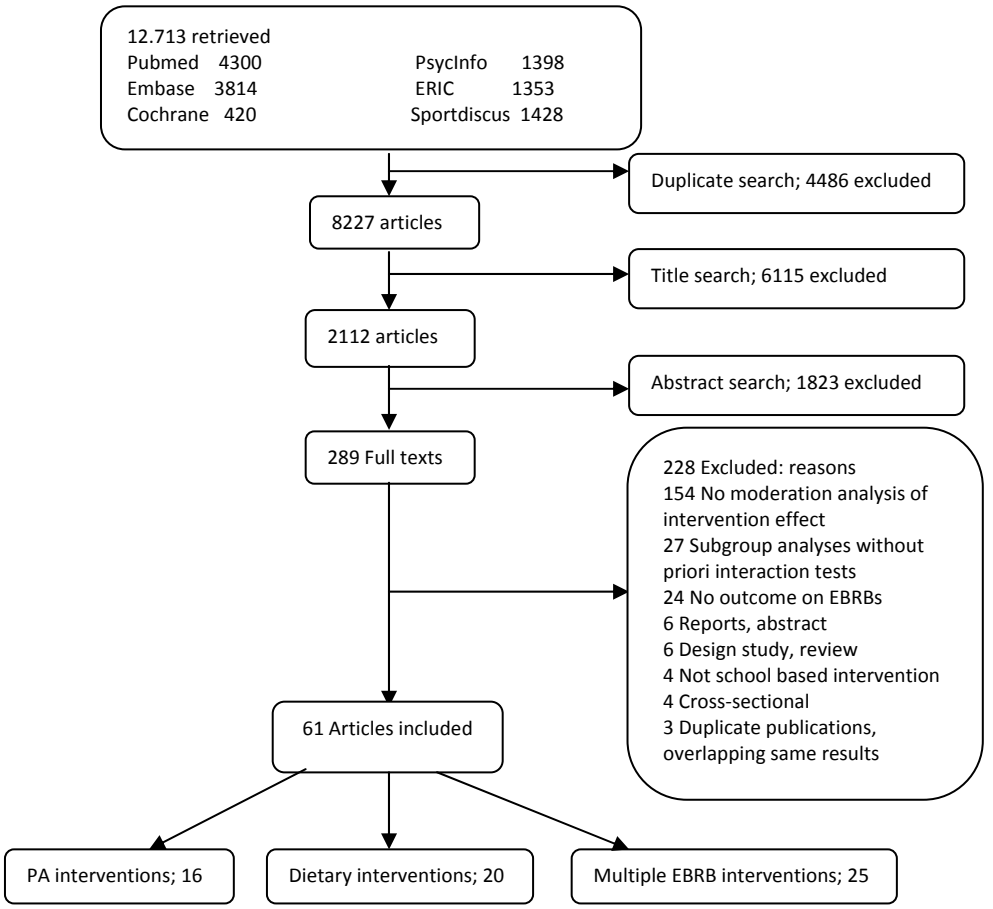

Figure 1. Article search and study selection

Description of interventions

Supplementary Table 3 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 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 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.

Table 1. The results of the moderation analyses of the school-based interventions aimed at energy balance-related behaviors (EBRBs).

EBRBs Variables	Total PA	Sedentary Behavior	Active Transport	Fat Intake	FV intake	Total Energy Intake	Snacking/ fast food	Sugar /Sweets Intake	Soft Drink Consumption
Personal									
Gender (female)	0,0,0,0,0,0,0, 0,0,0,0,0,0,0, 0,0,0,+,+,+,+, +,-	0,0,0,0,0,0, 0,-,-	0,0,0,+	0,0,0,0,0 ,0,0, 0,0,+,+,+	0,0,0,0, 0,0,0, 0,0,0,0,	0,0,0,0,0, 0	0,0	0,-	0,≈
Age (younger)	0,0,0,0	0,0		0,0,0,±	0,0,0,0, 0,0,0,	0	0		
Ethnicity (natives)	0,0,0,0,0,≈[a]	0,0,0,0,0	0	0,0, ≈(girls),≈	0,0,0,0, 0,0,0, 0,	0,0,0	0		0
Baseline values of outcome (unfavorable)	0,0,0,0,0,0,+			0,0,+	0,+,+	0,0	+		
Baseline weight status (obese)	0,0,0,0,0,0	+,0,0	0	0,0	0,0	0,0			
SES (low)	0,0,0,0,0,+	0	0	0,0,0	0,0,0,0,	0,0	+	0	0
Preferences					0				
Health status (baseline CVD risk (high))				0				+ (boys)	
Growth (increase in heights (low))	0			0		0			
Situational									
Recess time	0								
Site (State, province)	0,0	0		0	0,0	0			
Type of education (technical)				+[b]					
Lesson location (indoor)	+ VPA - walking								

EBRBs Variables	Total PA	Sedentary Behavior	Active Transport	Fat Intake	FV intake	Total Energy Intake	Snacking/ fast food	Sugar /Sweets Intake	Soft Drink Consumption
Situational									
Teacher speciality (classroom teacher)	+								
School food policy (fruit only policy)					+		0	0	
Compliance (high)	+								
Sports club participation	0	0	0						
Country					+				
Intervention part									
Family involvement (high)	0	0			+				
Food type provided (fruit)					0,≈,≈				
Psychosocial									
Intention (low)	+								
PBC (low)	+								
Attitude (low)	+								
Normative beliefs	0								

O ; Not significant, +; Significant for the group mentioned in brackets, -; Significant in the opposite direction for the group mentioned in brackets
x; Not significant after stratified analysis , **±** ; Stratified analysis results are not reported, **≈;** Significant for both groups
a; Significant in low income group, **b;** Significant in Natives, **VPA**; Vigorous physical activity, **FV**; Fruit and vegetable, **SES**; Socioeconomic status, **CVD**;Cardiovascular diseases, **PBC**; Perceived behavioral control

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.

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 metaregression 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 behaviors: 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 Jan 21; 6: 6.
8. Donaldson SI. Mediator and moderator analysis in program development. In S.Sussman (Ed.), Handbook of Program Development for Health Behavior Research and Practice (pp.470-500). Thousand Oaks, CA: Sage Publications, Inc. 2001.
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. First ed. New York: The Guilford Press; 2004.
12. Brookes ST, Whitely E, Egger M, Smith G.D., Mulheran PA, Peters TJ. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. J Clin Epid. 2004;57:229-36.
13. Cui L, Hung HMJ, Wang SJ, Tsong Y. Issues Related to Subgroup Analysis in Clinical Trials. J Biopharm Stat. 2002;12(3):347-58.
14. Jaccard J, Turrisi R, Wan CK. Interaction effects in multiple regression. First 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 Epid. 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. Public 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 Feb;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 Jan;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 Apr;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 Apr;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 Aug;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 Sep;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 Feb;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. Dzewaltowski 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 Medicine. 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 Dec;37(6 Pt 2):62-9.
36. Goran MI, Reynolds K. Interactive multimedia for promoting physical activity (IMPACT) in children. Obes Res. 2005 Apr;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 Nov;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 Jun;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. Public Health Nutr. 2007 May;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 Mar;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 Supplement):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 Sep;65(5):1049-58.
45. Himes JH, Ring K, Gittelsohn J, et al. Impact of the Pathways intervention on dietary intakes of American Indian schoolchildren. Prev Med. 2003 Dec;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 Dec;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 May;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 Mar;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 Jun;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 Mar;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 Mar 13;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. Public Health Nutr. 1999 Sep;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 Feb;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. Public Health Nutr. 2008 Mar;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 Aug;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 Community Health. 2008 Oct;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 Public Health. 1998 Apr;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 Nov;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 Feb;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? Public Health Nutr. 2007 Sep;10(9):939-47.
63. Robinson TN. Reducing children's television viewing to prevent obesity: a randomized controlled trial. JAMA. 1999 Oct 27;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 behaviors and promote physical activity in 10-year-old children: Switch-play. Int J Obes. (Lond) 2008 Apr;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 Oct;32(10):1489-98.
67. Singh AS, Chinapaw M, 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 Dec;37(6 Pt 2):70-79.
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):1059-1068.

70. Tak NI, Te Velde SJ, Brug J. Ethnic differences in 1-year follow-up effect of the Dutch Schoolgruiten Project - promoting fruit and vegetable consumption among primary-school children. Public Health Nutr. 2007 Dec;10(12):1497-507.
71. Tak NI, Te Velde SJ, Brug J. Long-term effects of the Dutch Schoolgruiten Project—promoting fruit and vegetable consumption among primary-school children. Public Health Nutr. 2009 Aug;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 Apr;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 Jan;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. JTPE. 2007 Jan;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 Public Health. 2006 Aug;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. Public Health Nutr. 2007 May;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 Mar;34(3):173-84.
78. Wechsler H, Basch CE, Zybert 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 Public Health. 1998 Mar;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 Dec;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 Epid. 2007;60:34-42.
81. Biddle SJH, Gorely T, Stensel DJ. Health-enhancing physical activity and sedentary behavior in children and adolescents. J Sport Scien. 2004;22:679-701.
82. Sweeting HN. Gendered dimensions of obesity in childhood and adolescence. Nutrition Journal. 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 Medical Research Methodology. 2008;8:47.
84. Simen-Kapeu A, Veugelers PJ. Should public health interventions aimed at reducing childhood overweight and obesity be gender-focused? BMC Public 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. Vilhjalmsson 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, Lowis 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 J. Environmental influences on energy balance-related behaviors: 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;19:617-622.
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 Methods. 2002;7(1):19-40.
95. Stone-Romero EF, Liakhovitski D. Strategies for detecting moderator variables: a review of conceptual and empirical issues. Research in Personnel and Human Resources Management. 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 Methods. 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 Intervention. 2007;29(3):262-72.

SUPPLEMENTARY FILES

Supplementary Table 1. Online search strategy for the database of Pubmed

 # 1 Children
"child"[MeSH] OR child* OR schoolchild* OR "adolescent"[MeSH] OR adolescen* OR boy OR boys OR girl OR girls OR youth OR youths OR teen OR teens OR teenager* OR puberty OR pupils
 # 2 Energy Balance-Related Behaviors
"physical activity" OR exercise OR exercising OR sport OR active OR inactive OR activities OR activity OR walking OR cycling OR sedentary OR computer OR television OR TV OR internet OR games OR gaming OR recreation OR play OR viewing OR nutrition OR food OR diet OR intake OR beverages OR snack OR "soft drink" OR softdrink OR fruit OR vegetable OR drink OR soda OR breakfast OR supper OR lunch OR meal OR dinner OR tea OR fat OR "5-a-day" OR eating OR sitting OR obese OR obesit* OR overweight OR "body mass" OR "body weight" OR "weight gain" OR "weight loss" OR BMI
 # 3 School-Based Intervention
(prevention OR intervention OR promotion OR trial OR program*) AND (school OR school* OR "school-based" OR education OR grade)
 # 4 Filter of unrelated studies
NOT infant NOT neonatal NOT alcohol NOT cigarette NOT tobacco NOT cannabis NOT marijuana NOT drug NOT cocaine NOT HIV NOT AIDS NOT condom NOT sexual NOT anorexia NOT bulimia NOT binge NOT autism NOT autistic NOT steroid NOT menopause NOT osteopenia NOT osteoporosis NOT menstruation NOT HRT NOT Schizophrenia NOT mental NOT depression NOT dementia NOT injury NOT HPV NOT vaccine NOT allergy NOT folate NOT folic NOT dental NOT caries NOT oral NOT teeth NOT sleep NOT apnea NOT asthma NOT tumour NOT cancer NOT chemotherapy NOT lymphoma NOT carcinoma NOT hepatitis NOT sun NOT dermatitis NOT eczema NOT acne NOT celiac NOT bullying NOT suicide NOT migraine NOT headache NOT kidney NOT infection NOT hyperactivity NOT cerebral NOT sickle NOT helicobacter NOT arthritis NOT pregnancy NOT contraceptive NOT fertility NOT dysmenorrhea NOT dysmenorrhoea NOT spina NOT fatigue NOT spondylitis NOT thalassemia NOT alzheimer NOT phenylketonuria NOT influenza NOT anaemia NOT iron NOT deficiency NOT malaria NOT measles NOT polio NOT traffic NOT helmet NOT immunization NOT vaccination NOT surgery NOT hygiene NOT gambling NOT fracture NOT stress NOT crohn
End Search # 1 AND # 2 AND # 3 AND #4

Supplementary Table 2. Quality assessment of the school-based interventions aimed at energy balance-related behaviors that conducted moderation analysis.

	Quality criteria / Study	Year	Randomization performed	Groups similar at baseline (regarding age, gender, BMI, outcome measures)	Eligibility criteria specified	Point estimates and variability measures of outcome	Intention-to-treat analyses included	Drop-out rate non selective	Theoretical framework cited	Did reliability and validity checked for outcome measure?	Was a rationale (evidence or theory) for the interaction provided?	Was power calculated and was it adequate to detect moderation?	Was the assumption of homogeneous (error) variances across groups checked?	**Total score**
1	Baranowski[20]	2000	+	+	+	+	+	+	+	+	-	-	-	8
2	Haerens[41]	2007	+	?	+	+	+	+	+	+	+	-	-	8
3	Young[79]	2006	+	+	+	+	+	?	+	+	+	-	-	8
4	Bere[24]	2006	+	+	+	+	-	+	+	+	-	-	?	7
5	Bere[23]	2006	+	+	+	+	-	+	+	+	-	-	?	7
6	Hill[44]	2007	+	+	+	+	+	+	+	-	-	-	-	7
7	Simon[66]	2008	+	-	+	+	+	-	+	+	-	-	+	7
8	Perry[60]	1998	+	+	+	+	-	+	+	+	-	-	-	7
9	Verstraete[74]	2007	+	+	+	+	-	+	-	+	+	-	-	7
10	Dzewaltowski[30]	2009	+	+	+	+	-	-	+	+	-	-	-	6
11	Goran[36]	2005	+	-	+	+	?	+	+	+	-	-	-	6
12	Haerens[38]	2007	+	+	+	+	-	?	+	+	-	-	-	6
13	Haerens[39]	2007	+	+	+	+	-	?	+	+	-	-	-	6
14	Himes[45]	2003	+	+	+	+	-	?	+	+	-	-	-	6
15	Kipping[49]	2008	+	+	-	+	+	?	-	+	+	-	-	6
16	Manios[54]	1999	+	?	+	+	-	+	+	+	-	-	?	6
17	Manios[53]	1999	+	?	+	+	-	+	+	+	-	-	?	6
18	Martens[56]	2007	+	?	+	+	+	-	+	+	-	-	-	6
19	Te Velde[72]	2008	+	-	+	+	+	-	+	+	-	-	-	6
20	Verstraete[76]	2007	+	+	+	+	-	+	-	+	-	-	-	6
21	Verstraete[75]	2006	+	+	+	+	-	+	-	+	-	-	-	6
22	Baranowski[21]	2003	+	-	+	+	-	-	+	+	-	-	-	5
23	Burke[26]	1996	+	+	+	-	-	+	-	-	+	-	-	5
24	Cardon[28]	2009	+	-	+	+	-	+	-	+	-	-	-	5
25	Cullen[29]	2005	+	-	+	+	?	-	+	+	-	-	-	5
26	Foster[32]	2008	+	-	+	-	+	+	-	+	-	-	-	5
27	Gentile[34]	2009	+	-	+	+	+	?	+	-	-	-	-	5

Supplementary Table 2 (continued)

28	Going[35]	2003	+	-	+	+	-	?	+	+	-	-	-	5
29	Haerens[40]	2007	+	?	+	+	-	?	+	+	-	-	-	5
30	Harrison[42]	2006	-	+	+	+	?	?	+	+	-	-	-	5
31	Hendy[43]	2007	+	+	+	-	-	?	-	+	+	-	-	5
32	Luepker[52]	1996	+	+	+	+	?	-	+	-	-	-	-	5
33	McKenzie[57]	2001	+	?	+	+	?	?	+	+	-	-	-	5
34	Moore[58]	2008	+	+	+	+	-	?	-	+	-	-	-	5
35	Perry[61]	2004	+	+	-	-	-	+	+	+	-	-	-	5
36	Robinson[63]	1999	+	-	-	+	+	?	+	+	-	-	-	5
37	Shilts[65]	2009	+	-	+	-	-	+	+	+	-	-	-	5
38	Singh[67]	2009	+	-	+	+	-	-	+	+	-	-	-	5
39	Stevens[68]	2003	+	?	+	-	+	?	+	+	-	-	-	5
40	Perry[59]	1998	+	+	-	-	-	+	+	+	-	-	-	5
41	Salmon[64]	2008	+	-	+	-	+	?	+	-	+	-	-	5
42	Webber[77]	2008	+	+	+	-	+	?	+	-	-	-	-	5
43	Ashfield-	2008	+	-	+	+	-	-	-	+	-	-	-	4
44	Barnett[22]	2009	+	?	+	-	-	-	+	+	-	-	-	4
45	Bere[25]	2005	+	-	+	+	-	-	-	+	-	-	?	4
46	Ernst[31]	1999	+	+	-	+	-	?	-	+	-	-	-	4
47	Graham[37]	2008	-	+	+	+	-	-	-	-	+	-	-	4
48	Kelder[48]	1995	-	+	+	-	-	?	+	+	-	-	-	4
49	Loucaides[50]	2009	+	-	+	+	-	?	-	+	-	-	-	4
50	Marcus[55]	2009	+	-	+	-	+	+	-	-	-	-	-	4
51	Reinaerts[62]	2007	+	-	+	+	-	-	+	?	-	-	-	4
52	Wechsler[78]	1998	+	+	+	-	-	?	-	+	-	-	-	4
53	Butcher[27]	2007	+	?	+	-	-	?	-	+	-	-	-	3
54	Horne[46]	2004	-	?	+	+	-	?	-	+	-	-	-	3
55	Horne[47]	2009	+	+	?	-	-	?	-	+	-	-	-	3
56	Stock[69]	2007	-	-	+	+	-	?	-	+	-	-	-	3
57	Tak[70]	2007	-	-	+	+	-	-	-	+	-	-	-	3
58	Tak[71]	2009	-	-	+	+	-	-	-	+	-	-	-	3
59	Vandongen[73]	1995	+	-	+	+	-	-	-	-	-	-	-	3
60	Frenn[33]	2003	-	-	-	-	-	?	+	+	-	-	-	2
61	Lowe[51]	2004	-	-	+	-	-	?	-	+	-	-	-	2
	TOTAL		52	27	54	44	15	19	37	51	8	0	1	Mean=5.1

Yes= +, No= - , Don't know=?

Supplementary Table 3. Characteristics of the included school-based interventions aimed at energy balance-related behaviors.

Study (Year) (Quality Score)	Sample	Design	Assessment	Theory	EBRBs Outcome (reliability and/or validity)	Intervention effect on outcome	Moderator tested	Results of moderator test	In case of moderators: results of stratified analyses
DIETARY INTERVENTIONS									
Ashfield-Watt[19] (2008) (low:4)	New Zealand 20 primary schools n=2032 %female=50.1 Mean age=8.8 Drop-out=44.0%	RCT I: 10 weeks, free fruit (at least one piece per school day) C: No intervention	Baseline; 10 and 16 weeks post-baseline	-	Fruit intake per school day by Day in the Life Questionnaire (DILQ) (kappa=0.85-0.92)	Significant intervention effect on fruit intake at 10 weeks	- Gender - Age - Ethnicity	NS NS NS	
Baranowski [20] (2000) (medium:8)	USA (Gimme 5) 16 elementary schools n=1732 %female=NR Mean age=NR Drop-out=32.3%	RCT I: 6 weeks/year (for 2 years), curriculum, parent involvement, videotapes and daily newsletters to home, point-of-purchase education C: No intervention	Baseline; 12 and 24 months post-baseline	SCT	FJV consumption by 7 day food record (daily servings) (r=0.80-0.98, ICC for F=0.72-0.76, ICC for V=0.59-0.69)	Significant intervention effect on combined FV intake and vegetables alone at the 1st year	- Gender - Ethnicity	NS NS	
Baranowski [21] (2003) (medium:5)	USA (Squire's Quest!) 26 elementary schools n=1578 %female=50.9 Mean age= 8.3 Drop-out=5.3%	RCT I: 5 weeks, Interactive multimedia game, goal setting C: No intervention	Baseline; 5 weeks post-baseline	SCT	FJV consumption by 4 non-consecutive days food intake recording software system (r=0.44-0.59)	Significant intervention effect on fruit intake, regular vegetables intake and total FJV intake	- Gender - Age - Ethnicity	NS NS NS	
Bere[24] (2006) (medium:7)	Norway (Fruit and Vegetables Make the Marks-FVMM) 19 schools (9 intervention) n=450 %female=54.2 Mean age=11.3 Dropout=31.0%	RCT I: 7 months, curriculum, parental involvement, paid school fruit programme (only at 4 schools) C: 3 schools in paid school fruit programme, others no intervention	Baseline; 8 and 20 months post-baseline	SCT	FV consumption by 24 hour recall (test-retest r=0.62-0.83, validity r=0.21-0.32)	NS	- Gender	NS	

Study (Year) (Quality Score)	Sample	Design	Assessment	Theory	EBRBs Outcome (reliability and/or validity)	Intervention effect on outcome	Moderator tested	Results of moderator test	In case of moderators: results of stratified analyses
Bere[23] (2006) (medium:7)	Norway (Norwegian School Fruit Programme + FVMM) 19 schools (9 intervention) n=577, %female=47.6 Mean age=11.3 Drop-out=10.4%	RCT I: 7 months curriculum, parental involvement, school fruit programme (1st year free, 2nd year paid fruit only in 4 schools) C: 1st year 2, 2nd year 3 schools in school fruit programme, others no intervention	Baseline; 8 months and 20 months post-baseline	SCT	FV consumption by 24 hour recall and food frequency questionnaire (test re-test r=0.62-0.83, validity r=0.21-0.32)	Significant intervention effect on FV intake in both follow-ups	- Gender - Habitual FV intake - Preferences of FV - SES	NS in Year 1, Significant in Year 2 NS NS NS	Stratified by gender: Boys > Girls (NS)
Bere[25] (2005) (low:4)	Norway (Norwegian School Fruit Programme) 38 schools (9 free fruit, 9 paid fruit, 20 control) n= 922, %female=50.1 Mean age=12.3 Drop-out=13.8%	3-arm RCT I:1- free fruit programme, 2- paid fruit programme C: No intervention	Baseline; 8 months post-baseline	-	FV consumption by 24 hour recall and food frequency questionnaire (test re-test r=0.62-0.83, validity r=0.21-0.32)	Significant intervention effect among free fruit group compared to paid and no fruit group Significant intervention effect among subscribers at the paid fruit group compared to non-subscribers	- Gender - Habitual FV intake - SES	NS Significant in FV intake at school and snacks Significant in soda/candy/chips intake	Stratified by habitual FV intake; Low habit > high habit Stratified by SES: Low SES > high SES
Cullen[29] (2005) (medium:5)	USA (Squire's Quest!) 26 elementary schools (13 intervention) n=1578, %female=50.9 Mean age= 8.3 Drop-out=5.3%	RCT I: 5 weeks Interactive multimedia game, goal setting C: No intervention	Baseline; 5 weeks post-baseline	SCT	FJV consumption at specific meals and snacks by 4 days food intake recording software system (r=0.44-0.59)	Significant intervention effect on servings of fruit and 100% fruit juice at snacks, and regular vegetables at lunch	- Gender - Age - Ethnicity - SES	NS NS NS NS	
Hendy[43] (2007) (medium=8)	USA (Kids choice programme) 1 school n=313, %female=46.2	3-arm RCT I:1) Changes in school lunch procedures, token reinforcement for vegetables,	Baseline; 2 weeks and 7 months post-	-	FV consumption by lunch observations, 3 days a week (inter-observer r= 0.92 for fruit consumption and	Significant intervention effect on FV intake at both follow-ups in	-Weight status -Food type provided	Study phase x Food type provided x Weight; NS	

	Mean age=8.0 Drop-out=12.1%	2) Changes in school lunch procedures, token reinforcement for fruits C: Changes in school lunch procedures	baseline		r=0.94 for vegetable consumption)	both reinforcement conditions			
Horne[46] (2004) (low=3)	UK (The Food Dudes) 2 primary schools n=749, %female=NR Age range=5-11 Drop-out=NR	Quasi experimental I;16 days, peer modelling videos, rewards, home packs, letters, free FV availability at school lunch for 24 days, also at snack time for 5-7 y olds, 4 months maintenance without videos C:No Food Dudes program but free availability of FV at lunch and snack time	Baseline; 1 month and 5 months post-baseline	-	FV consumption by lunchtime observation (kappa coefficient=0.89-0.96) snack time consumption by weighed FV before and after consumption (r=NR) and home consumption by parental recall (r=NR)	Significant intervention effect on lunchtime, snack time and home FV consumption at both follow-ups	-Free food type (free fruit, free vegetable) - Age (5-7y, 7-11y)	Significant on lunchtime FV consumption <u>School x Study phase x Food x Age</u>; Significant on lunchtime FV consumption	<u>Stratified by food type and age group;</u> Fruit and Vegetable both significant Younger > Older
Horne[47] (2009) (low=3)	Ireland (The Food Dudes) 2 primary schools n=435, %female=NR, Age range=4-11 Drop-out=NR	RCT I;16 days, peer modelling videos, rewards, homepack, letters, free FV availability at school lunch, maintenance phase without videos, supporting parents for providing FV in lunchboxes, rewards C:No intervention	Baseline; 12 months post-baseline	-	Consumption of FV at lunchboxes by Lunchbox Measures 5 days (r=NR) Consumption of free FV provided by school by School-provided foods measures 4 days (r=NR)	Significant on both consumption of lunchboxes FJV and school provided FV	- Food type (fruit, vegetable) - Age (4-7y, 7-11y)	Time x Group x Food ; Significant for total consumption of school-provided FV NS	<u>Stratified by food</u>; Fruit and Vegetable both significant
Lowe[51] (2004) (low=2)	UK (The Food Dudes) 3 primary schools n=402, %female=NR, Age range=4-11 Drop-out=NR	Quasi experimental I;16 days, 1) peer modelling videos, rewards, letters, free FV availability at school lunch at snack time, 2) 16 days, peer modelling videos, rewards, letters, free FV availability at school lunch at snack time, homepacks C:No intervention	Baseline; post interventi on (Time interval was not reported)	-	FV consumption by lunchtime observation (kappa coefficient=0.96-0.99) snack time consumption by weighed measure (r=NR) and home consumption by parental recall (r=NR)	Significant intervention effects on lunchtime, snack time consumption and weekdays consumption outside school.	- Age (4-7y, 7-11y)	Significant on lunchtime fruit consumption	<u>Stratified by age;</u> Younger > Older

Study (Year) (Quality Score)	Sample	Design	Assessment	Theory	EBRBs Outcome (reliability and/or validity)	Intervention effect on outcome	Moderator tested	Results of moderator test	In case of moderators: results of stratified analyses
Martens[56] (2007) (medium=6)	The Netherlands (Krachtvoer) 18 schools n=1613 %female=NR Age range=12-14 Drop-out=22.5%	RCT I:curriculum, materials (lunch-boxes, take home bag, magazine, posters, postcards, video, website, newsletters, taste-testing) C: No intervention	Baseline; 3 months post-baseline	ASE SM	FJ consumption by a self-administered 24-hour recall (r=0.53-0.64) Snack consumption by a short food frequency questionnaire, the Fat list (r=0.60) Breakfast consumption by 1 item	Significant effect on fruit consumption	- Gender - Age - Baseline value (centered- low, medium, high)	Significant on fruit frequency Significant on total fat consumption at breakfast Significant on all outcome variables	Stratified by gender and age; NR Stratified by baseline value; - Fruit consumption Low, medium > high - Fruit juice consumption; High > Low -Snack consumption; High > Low -Breakfast saturated fat intake; High > Low
Moore[58] (2008) (medium=5)	UK 43 primary schools n=1632 %female=51.3 Age range=9-11 Drop-out=NR	RCT I:1 school year, fruit tuck shops at schools C: No intervention	Baseline; 9 months post-baseline	-	Fruit, sweet (sweets, chocolate, biscuits) and savory snacks (crisps) consumption at school and 24 h by a single-day computerized 24-hour recall questionnaire (kappa=0.29)	Significant effect on fruit intake as snacks at school	-School food policy	Significant on fruit consumed at school	Stratified by school policy; Fruit only policy > No policy
Perry[61] (2004) (medium=5)	USA (5-a-day Cafeteria Power Plus) 26 schools n=1668 % female=49 Mean age=NR Drop-out=30%	RCT I: 2 school year, school food service change, social support C: No intervention	Baseline; 2 years post-baseline	SCT HBPM	FV consumption at lunch by lunch observation (r=NR)	Significant intervention effect on total FV intake	- Gender - Age	NS NS	
Perry[59] (1998) (medium=5)	USA (5-a-day Power Plus) 20 elementary schools n=536 % female=50.0 Mean age=10.0 Drop-out=17.7%	RCT I: Behavioral curricula, parental involvement, school food service changes, industry support C: No intervention	Baseline; 12 months post-baseline	SLT	FV consumption and total dietary intake (total energy, fat, saturated fat) by 24-h recall (r=0.45-0.79) and lunch observation (r=NR)	Significant intervention effect on FV consumption	- Gender - Ethnicity	Significant on vegetable consumption Significant on total fat and saturated fat intake	Stratified by gender; Girls > Boys Stratified by ethnicity; Asians, African-Americans > Hispanics, Caucasian
Reinaerts[62] (2007)	The Netherlands 12 primary schools	3-arm RCT I: 9 months I1) Free	Baseline; 9 months	SCT ELM	FV consumption by pre-structured 24-h food	Both intervention	-Gender	Significant on vegetable intake	Stratified by gender; Girls > Boys

(low=4)	n=1730 % female=51.5 Mean age=9.0 Drop-out=45.7%	FV distribution (twice a week), I2) Classroom curriculum, parental involvement C: No intervention	post-baseline	HBM HT	recall, FFQ - parents reported (r=NR)	arms had significant effect on FV consumption	-Ethnicity -Age	Significant on vegetable intake and 24 h FJV intake Significant on vegetable intake and 24 h FJV intake	<u>Stratified by ethnicity;</u> Non-Native > Native <u>Stratified by age;</u> Older > Younger
Tak[70] (2007) (low=3)	The Netherlands (The Dutch SchoolGruiten Project) 55 schools n= 1328 % female=55 Mean age=10.0 Drop-out= 28.2%	Quasi experimental I: 2 years, Free piece of fruit or vegetable (twice a week), school curriculum to increase knowledge and skill related to F&V consumption (not obligatory) C: No intervention	Baseline; 1 year post-baseline	-	FV intake by FFQ both from children (r=0.47-0.84) and parents (validity r=0.40-0.53, test-retest r=0.47-0.84)	Significant intervention effect on vegetable intake of non-Western children Significant intervention effect on fruit intake of Dutch children	-Gender -SES (low, medium, high)	Significant on vegetable intake of Dutch children Significant on fruit intake of Non-western children	<u>Stratified by gender;</u> Girls > Boys <u>Stratified by SES;</u> NS difference
Tak[71] (2009) (low=3)	The Netherlands (The Dutch SchoolGruiten Project) 55 schools n=1328 % female=55.1 Mean age=10.0 Drop-out: 41.9%	Quasi experimental I: 2 years, Free piece of fruit or vegetable (twice a week), school curriculum to increase knowledge and skill related to F&V consumption (not obligatory) C: No intervention	Baseline; 1 year and 2 years post-baseline	-	FV intake by FFQ both from children (r=0.47-0.84) and parents (validity r=0.40-0.53, test-retest r=0.47-0.84)	Significant increase in fruit intake at the 2nd follow up both in children and parent reports No significant effect on vegetable intake	-Gender -Ethnicity -SES -Region of residence	NS NS Significant on child-reported fruit intake NS	<u>Stratified by SES;</u>NS difference
Te Velde[72] (2008) (medium=6)	Norway, The Netherlands, Spain (Prochildren) 62 schools n=1801 %female=61.2 Age range=10-13 Drop-out=23.8%	RCT I: 2 school years, classroom component (activity worksheets and computer-tailored feedback), school component (free or paid FV), family component (homework assignments, newsletters, parent version of computer-tailored tool), optional component (community participation) C: No intervention	Baseline; 20 and 32 months post-baseline	SCT ASE TTI SEF	Total and separate FV intake by 24 h recall, and by FFQ (validity r=0.40-0.53, test-retest r=0.47-0.84)	Significant effect on total and separate FV intake at the 1st follow-up	-Gender -Country	NS Significant at 2nd follow-up at total FV intake and total fruit intake	<u>Stratified by country</u> Norway > Netherlands, Spain

Study (Year) (Quality Score)	Sample	Design	Assessment	Theory	EBRBs Outcome (reliability and/or validity)	Intervention effect on outcome	Moderator tested	Results of moderator test	In case of moderators: results of stratified analyses
Wechsler[78] (1998) (low=4)	USA 6 elementary schools n=6902 %female=NR Mean age=NR Drop-out= None (Unit of analysis was school)	RCT I:7-10 days, social marketing techniques (e.g., taste tests, promotion incentives, product positioning), auditorium session, flyers C: No intervention	Baseline; 7 to 8 days, 3 to 4 months post-baseline	-	Lunchtime low-fat milk consumption by sampling discarded milk cartons (inter-rater r=0.99)	NS on milk consumption	-Gender -Age (younger (1^{st}-2^{nd} grades) vs. older (3^{rd}-4^{th} grade))	NS NS on consumption	
PHYSICAL ACTIVITY INTERVENTIONS									
Barnett[22] (2009) (low:4)	Australia (The Physical Activity and Skills Study) 18 schools n=928 %female=53.0 Mean age=10.1 Drop-out=70.3%	Quasi-experimental I: 1 year, Modification of existing PE classes C:No intervention	Baseline; 1 year and 6 years post-baseline	-	PA in a week by the Adolescent Activity Recall Questionnaire (APARQ) (r=0.35-0.88, kappa=0.25-0.74)	NS	- Gender - Age	NS NS	
Butcher[27] (2007) (low=3)	UK 3 schools (1 per condition) n=177, %female= 45.8 Age range= 7-11 Drop-out=20.3%	3-arm RCT I:1- Feedback on pedometer steps 2- Feedback and information on pedometer steps C: No intervention	Baseline; 1 week post-baseline	-	Daily step counts by pedometer (ICC=0.90-0.94)	Significant intervention effect on daily step counts in feedback + information group compared to other groups.	- Gender	NS	
Cardon[28] (2009) (low=4)	Belgium 40 pre-schools (10 per condition) n=636, %female=47.8 Mean age=5.3 Drop-out= 8.3%	RCT I:1-play equipment provided, 2- painted markings on the playground, 3-play equipment + painted markings on the playground C:No intervention	Baseline; 3 months post-baseline	-	PA level during recess by accelerometers	NS	- Gender - Age - Pre-test values - Average recess time	NS NS NS NS	
Ernst[31] (1999) (low=4)	USA (Promoting Lifetime Activity for Youth-PLAY) 5 elementary	RCT I:12 weeks, 15 min PA break each school day with classroom	Baseline; 4 weeks and 12 weeks	-	PA level by 7 days PA Questionnaire for Older Children (PAQ-C) (test-retest r=0.75-	Significant intervention effect on PA level in	-Baseline activity level	NS	

	schools n=NR %female=NR Mean age=NR Drop-out=NR	teachers support, goal setting and recording handbooks C: 15 min activity break without teachers support, recording handbooks	post baseline		0.82) (concurrent validity r=0.45-0.53)	intervention group			
Goran[36] (2005) (medium=6)	USA (IMPACT) 4 elementary schools n=209, %female= 51.0 Mean age=9.5 Drop-out=40.7%	RCT I: 8 weeks, Interactive educational learning game C:Popular educational CDs (unrelated to health)	Baseline; 2 and 4 months post-baseline	SCT	PA by 5 days uniaxis accelerometer (r=0.80-0.92)	NS on total PA Significant effects on MVPA and LPA	- Gender	Significant on LPA	<u>Stratified by gender;</u> Girls > Boys
Graham[37] (2008) (low=4)	USA 2 high schools n=122 %female=100 Mean age=15.0 Drop-out:2.5%	Quasi experimental I: 1 school year intervention: 4 weekly supervised exercise sessions + 1 weekly health education C: no intervention	Baseline; 9 months post-baseline	-	PA- from 3 day PA recall aggregated into moderate (MPA) and vigorous (VPA) activities (r=NR)	Significant intervention effects on VPA	-TV viewing (high vs. low)	Significant on VPA	<u>Stratified by TV viewing;</u> High > Low
Harrison[42] (2006) (medium=5)	Ireland (Switch off-Get Active) 9 primary schools n= 312, %female=43.3 Mean age=10.2 Drop-out=9%	Quasi experimental I: 16 weeks, health education, diaries signed by parents, homeworks C: No intervention	Baseline; 16 weeks post-baseline	SCT	PA and screen time by 1-day Previous day PA recall (test-retest r=0.98, validity against accelerometer r=0.88)	Significant intervention effect on physical activity but not on screen time	-Gender - BMI	NS NS	
Hill[44] (2007) (medium=7)	UK 1 secondary school n=620 51% girls Mean age: 16.9 Drop-out: 16%	4-arm RCT I: 1) Leaflet; 2) Leaflet + motivational incentive and quiz; 3) Leaflet + implementation intention prompt C: No intervention	Baseline; 3 weeks post-baseline	TPB ELM	Exercise- from 1 item weekly exercise measure (r=NR)	Significant intervention effects on weekly exercise. No differences between interventions	- Intention - PBC - Attitude - Normative beliefs	Significant Significant Significant NS	<u>Stratified analysis;</u> <u>Lower > Higher</u>
Loucaides[50] (2009) (low=4)	Cyprus 3 schools n= 247 %female=NR Mean age=11.1 Drop-out=7.7% (break time), 6.1% (after school)	RCT I:1) play space for team games, playground markings, providing equipment 2) play space for team games, providing equipment C: No intervention	Baseline; 4 weeks post-baseline	-	PA during school breaks and after school period by pedometer (r=0.53-0.81)	Significant intervention effect on break-time activity in both intervention groups compared to control school	- Gender	NS	

Study (Year) (Quality Score)	Sample	Design	Assessment	Theory	EBRBs Outcome (reliability and/or validity)	Intervention effect on outcome	Moderator tested	Results of moderator test	In case of moderators: results of stratified analyses
Salmon[64] (2008) (medium=5)	Australia (Switch-Play) 3 schools n=295 %female=51.0 Mean age=10.7 Drop-out=12%	4-arm RCT I: 9 months, I1) behavioral modification (BM) group; classroom based, parent involvement, I2) Fundamental movement skills (FMS) group; funny games for 6 movement skills, I3) BM+FMS C: No intervention	Baseline; **9** and 23 months post-baseline	SCT BCT	Physical activity by accelerometers (r=NR) Self-reported screen behaviors by a questionnaire (ICC>0.6)	Significant intervention effect on physical activity in BM and FMS intervention arms at both time points Significant intervention effect of BM on TV viewing in undesired direction at both timepoints	-Gender	Significant on movement counts, MPA and VPA and screen behaviors	Stratified by gender In FMS and BM+FMS groups Boys > Girls
Simon[66] (2008) (medium=7)	France (ICAPS) 8 middle schools n=954 %female=50.0 Mean age=11.6 Drop-out=23.3%	RCT I:4 years, educational classes, debates, opportunities for PA at lunchtime and breaks, sporting events, parental support C:No intervention	Baseline; 8, 20 and 32 months post-baseline	SEF	Self-reported leisure PA and TV/video viewing and active commuting by the Modifiable Activity Questionnaire (ICC=0.71-0.83)	Significant intervention effect on leisure time PA and at all time points TV/video viewing only at 20 months follow-up	-Gender -SES -Initial weight -Sports club participation	NS NS NS NS	
Verstraete [76] (2007) (medium=6)	Belgium 16 elementary schools n=810 %female=50.7 Mean age=9.7 Drop-out= 5.7%	RCT I:2 school years, Health-related PE programme, classroom-based health education lessons, an extra-curricular PA promotion programme C: No intervention	Baseline; 20 months post-baseline	-	PA level in leisure time by a PA questionnaire (ICC=0.68-0.93, validity r=0.27-0.44) PA level by a 5 day uniaxis accelerometer (validity r = 0.50-0.74)	Significant effect on MPA and MVPA, and on MPA in leisure time	-Gender	NS	
Verstraete [75] (2006) (medium=6)	Belgium 7 elementary schools n=249 %female=48.5 Mean age=10.8 Drop-out= 5.6%	RCT I:2 school years, Health-related PE programme, classroom-based health education	Baseline; 3 months post-baseline	-	PA level by uniaxis accelerometers at 1 day morning recess and lunch break (validity r = 0.50-0.74)	Significant effect on MPA and VPA at lunch break, on MPA at morning break	-Gender -Baseline MVPA	Significant on LPA, MPA and MVPA NS	Stratified by gender; Girls > Boys

		lessons, an extra-curricular PA promotion programme C: No intervention						
Verstraete [74] (2007) (medium=7)	Belgium 16 elementary schools n=791 %female=52.2 Mean age=9.7 Drop-out=3.4%	RCT I:2 school years, curricula, PE lessons, health-fitness and skill-fitness activities C: No intervention	Baseline; 20 months post-baseline	-	PA by accelerometer by uniaxis accelerometers (r=NR) PA in PE lessons by SOFIT (intrarater r=0.80-0.99, ICC= >0.95)	Significant effect on PA in PE lessons	-Gender	NS
Webber (2008) (medium=5)	USA (Trial of Activity for Adolescent Girls (TAAG)) 36 middle schools n= 1603(baseline)-3378 (follow-up) %female=100 Mean age=12 Drop-out=5.5% (Analysis unit was school)	RCT I: 2 years, environmental and organizational changes supportive of PA, cues, messages and incentives to be more PA, lessons to enhance behavioral skills C: No intervention	Baseline; 2 and 3 years post-baseline	OLT SCT OCT DIM	PA and sedentary behavior by 7 days wear of uniaxis accelerometers (r=NR)	Significant intervention effect on MVPA and sedentary behavior at the second follow-up	-Ethnicity	NS
Young [79] (2006) (medium=7)	USA 1 high school n=221 %female=100 Mean age= 13.8 Drop-out=5%	RCT I: 8 months, curricula, small-group discussion, homeworks, PE curriculum, weekly exercise logs and feedback on it, parental involvement C: Standard PE, newsletters to parents about PA content and general health interests	Baseline; 8 months post-baseline	SAT	PA by self-reported 7 day PA recall (test-retest r=0.59-0.81, validity r=0.76) Sedentary activities (TV viewing, computer, internet using) (test-retest r=0.94)	Significant effect on TV viewing during school days	-Ethnicity	NS
MULTIPLE EBRB INTERVENTIONS								
Burke [26] (1996) (medium:6)	Australia 30 primary schools n=1147 %female=43.8 Age range=10-12 Drop-out=15.3%	6-arm RCT I:9 months with 1- Physical fitness, 2-physical fitness+school nutrition, 3-school nutrition, 4-school + home nutrition, 5-home nutrition C:No intervention	Baseline; 10 months post-baseline	-	Nutrient intake (fibre, fat and sugar intake) by 2 days dietary recall (r=NR)	Significant intervention effect on nutrient intake only in RCT arms which included home nutrition.	- Cluster membership (High vs. Low cardiovascular risk)	Significant on sugar intake	<u>Stratified by cluster membership:</u> high risk boys > low risk boys in school+ home nutrition group

Study (Year) (Quality Score)	Sample	Design	Assessment	Theory	EBRBs Outcome (reliability and/or validity)	Intervention effect on outcome	Moderator tested	Results of moderator test	In case of moderators: results of stratified analyses
Dzewaltowski[30] (2009) (medium:6)	USA (Healthy Youth Places) 16 middle schools n=2211 %female=53.0 Mean age: 12.4 Drop-out: 28%	RCT I: intervention promoting personal and proxy agency to built healthy school environments that promote FV intake and PA C: No intervention	Baseline; 12 and 24 months post-baseline	SCT	PA- from previous day PA recall questionnaire (test-retest r=0.98, validity r=0.88) FV: 24-h Youth/ Adolescent FFQ (r=0.21-0.58)	Significant effect on VPA and MVPA NS on FV intake	- Gender - Ethnicity - SES - BMI	NS on PA NS on PA NS on PA ? Results on FV intake not reported	
Foster[32] (2008) (medium:5)	USA (School Nutrition Policy Initiative) 10 schools n=1349 %female=53.7 Mean age= 11.2 Drop-out=37.4%	RCT I: school self-assessment, nutrition education, nutrition policy, social marketing, parent outreach C: No intervention	Baseline; 24 months post baseline	-	Dietary intake (total energy, total fat, FV intake) by Youth/Adolescent Questionnaire (validity r=0.54, test -retest r=0.21-0.58), PA and sedentary behavior by Youth/Adolescent Activity Questionnaire (r=NR)	NS on dietary intake and PA Significant intervention effect on total inactivity and television watching	- Gender - Age - Ethnicity	NS NS NS	
Frenn[33] (2003) (low=1)	USA 2 middle schools n=341 %female=NR Age range=12-15 Drop-out=61.9%	Quasi-experimental I: 4 internet and video sessions, a healthy snack and gym labs C:No intervention	Baseline; post baseline (Time interval was not reported)	TM, HPM	Fat intake by Food habits questionnaire (α=0.82) Physical activity by Child adolescent activity log (α=0.77)	NS	- Gender - Ethnicity - SES	Gender x Group; Significant on fat intake Gender x Ethnicity x Group; Significant on fat intake Ethnicity x SES x group; Significant on MVPA	Stratified by gender and ethnicity; In girls, Asians < other ethnicities Stratified by ethnicity and SES ; In the low SES group; Native Americans < other ethnicities
Gentile[34] (2009) (medium=5)	USA (The Switch) 10 Elementary schools n=1196 %female=53.2 Mean age=9.6 Drop-out= 7.2%	RCT I: Community components as advertising key messages, school and family component as providing materials and goal setting	Baseline; 7 and 13 months post baseline	SEF	PA by pedometer Screen time (parent and child reported) FV consumption by items from National Youth Risk Behavior Survey (parent and child reported)	NS on PA level Significant intervention effect on parent reported screen time Significant intervention	- Gender - Family involvement - Weight status	Significant on FV intake and PA level Significant on FV intake Significant on screen time	Stratified by Gender: Girls > Boys Stratified by family involvement: High involvement > low involvement Stratified by weight status; Obese > normal

		C:No intervention (possible to expose the community component)			(r=NR)	effect on parent reported FV consumption			
Going[35] (2003) (medium=5)	USA (Pathways) 41 schools n=580 %female= 48.3 Mean age=7.6 Drop-out=52.1%	RCT I:3 year trial with a school curriculum, a PA/PE component, a school food service component, and a family component C:No intervention	Baseline; 3 years post-baseline	SLT	PA by 1 day triaxis accelerometer wearing (r=0.71)	NS	- Site	NS	
Haerens[38] (2007) (medium=6)	Belgium 10 secondary schools n=304 %female=70.4 Mean age=13.2 Drop-out=NR	RCT I: interactive computer tailored intervention which aimed dietary fat intake, tailored feedback C:No intervention	Baseline; 13 weeks post-baseline	TM TPB SCT ASE	Dietary fat intake by FFQ (α=0.83, validity r=0.78 against 7-day food diary)	Significant intervention effect on fat intake	- Gender - Type of Education	Condition x Gender x Education; Significant	Stratified analysis by gender and education; In girls; technical-vocational school > general schools
Haerens[41] (2007) (medium=8)	Belgium 15 schools n=2840 %female=36.6 Mean age=13.1 Drop-out=14.3%	3- arm RCT I: 1 school year intervention with 1) individual component: computer tailored PA intervention + environmental component: promoting PA availability, 2) components of intervention arm 1+ parental involvement C: No intervention	Baseline; 9 months post-baseline	TPB TM SCT	Total and leisure time PA level by Flemish PA Questionnaire (test-retest r>0.70, validity r=0.43-0.79) PA level by accelerometers (r=0.78-0.80)	Significant intervention effect on self-reported school-related PA Significant intervention effect on LPA and MVPA measured with accelerometers	- Gender	Significant for leisure time active transportation NS for accelerometer data	Stratified by gender; Girls > Boys
Haerens[39] (2007) (medium=6)	Belgium 10 secondary schools n=281 %female=48.8 Mean age=13.2 Drop-out=?	RCT I: interactive computer tailored intervention which aimed PA, tailored feedback C:No intervention	Baseline; 13 weeks post-baseline	TM TPB SCT ASE	PA level by Flemish PA Questionnaire (test-retest r>0.70, validity against accelerometer r=0.43-0.79)	Significant intervention effect on school related physical activity levels	- Gender -Compliance with the PA guidelines	NS NS	

Study (Year) (Quality Score)	Sample	Design	Assessment	Theory	EBRBs Outcome (reliability and/or validity)	Intervention effect on outcome	Moderator tested	Results of moderator test	In case of moderators: results of stratified analyses
Haerens[40] (2007) (medium=5)	Belgium 15 schools n=2840 %female=36.6 Mean age=13.1 Drop-out=15.7%	3- arm RCT I: 1 school year interventions with 1) individual component: computer tailored intervention on fat and fruit intake+ environmental component: promoting availability of healthy products 2) components of intervention arm 1 + parental involvement C: no intervention	Baseline; 9 months post-baseline	TPB TM SCT	Dietary fat intake by self-administered questionnaire (r=0.70-0.87, validity r=0.67-0.60 against 7-day food diary) Fruit intake by FFQ (r=0.52-0.82) Soft drink and water consumption by another FFQ (r=NR)	NS	- Gender	Significant on daily fat intake and percentage of energy from fat	Stratified by gender: Girls > Boys in parental support group compared with both other conditions.
Himes[45] (2003) (medium=5)	USA (Pathways) 41 schools n= 470 for lunch observation (n=620 for 24-h dietary recall) %female=48.3 Mean age=7.5 Drop-out=NR	RCT I:3 year trial with a school curriculum, a PA/PE component, a school food service component, and a family component C:No intervention	Baseline; 3 years post-baseline	SLT	Nutrients intake (total energy, fat, protein, carbohydrate, saturated and polyunsaturated fat, sucrose, fructose, fibre intake) by school lunch observation (r=0.96) and by single 24-hour dietary recall (r=NR)	Significant intervention effects on % of energy from total fat, saturated fat and total carbohydrate by lunch observation and 24-h recall. Significant effects on total energy, protein, total fat, saturated fat and polyunsaturated fat intake by 24-h recall	- Gender	NS	
Kelder[48] (1995) (low=4)	USA (The Class of 1989 Study) n=2376 %female=50.0 Mean age=NR	Quasi experimental I: 5 year, peer leaded education program, lunch bags, preparing own newspaper,	Baseline; 1, 2, 3, 4, 5 and 6 years post-	SCT	PA by self reported hours of exercise per week outside of gym class (test retest r=0.61)	? Results for the total study population were not reported	-Gender	NS except 3rd follow-up	Stratified by gender: Girls > Boys

	Drop-out=NR	monitoring aerobic activity, food preparation, community component; mass media and adult education, food labelling C: No intervention	baseline					
Kipping[49] (2008) (medium=6)	UK (Active for Life Year 5) 19 primary schools n= 604 %female=42.9 Mean age=9.41 Drop-out=23.6% (on activity data), 24.4% (on screen time data)	RCT I:5 months, PA lessons, nutrition lessons, screen viewing lesson C:No intervention	Baseline; 5 months post baseline	-	Time spent doing screen-based activities by a questionnaire (r=0.94) Active transport by 'A Day in the Life' questionnaire (r=NR)	NS	- Gender	NS
Luepker[52] (1996) (medium=5)	USA (CATCH) 96 elementary schools n=5106 %female= 48.2 Mean age= 8.8 Drop-out=21%	RCT I:3 years, classroom curricula, food service changes, physical education modifications, family participation C: No intervention	Baseline; 1, 2 and 3 years post-baseline	SCT	Total energy, fat, saturated fat, cholesterol intake by a 24 hour dietary recall (r=0.45-0.79) Leisure time PA, TV watching, video gaming by Self-administered physical activity checklist (validity r=0.60-0.76)	Significant effects on percentage of energy intake from fat, protein and carbohydrates, daily saturated fat, polyunsaturated and monounsaturated fat and cholesterol intake and on VPA	- Gender - Ethnicity - Site	NS NS NS
Manios[54] (1999) (medium=5)	Greece 21 primary schools n=579 %female=NR Mean age= NR Drop-out=18.7%	RCT I:3 years, curriculum, PE sessions (2/per week), parental involvement C: No intervention	Baseline; 3 years post-baseline	SLT	3 day physical activity measurement (out of school) by parent reported questionnaire (validity r = .68, test-retest r=0.64) Total energy, fat, protein, carbohydrates, fiber, cholesterol intake by 3 day weighed food record-parent reported (r=NR)	Significant effect on leisure time MVPA	-Gender -Baseline values -Parental education level -Increase in heights -Initial BMI	NS NS NS NS NS

Study (Year) (Quality Score)	Sample	Design	Assessment	Theory	EBRBs Outcome (reliability and/or validity)	Intervention effect on outcome	Moderator tested	Results of moderator test	In case of moderators: results of stratified analyses
Manios[53] (1999) (medium=5)	Greece 21 primary schools n=579 %female=NR Mean age= NR Drop-out=14.8%	RCT I:3 years, curriculum, PE sessions (2/per week), parental involvement C: No intervention	Baseline; 6 years post-baseline	SLT	3 day physical activity measurement (out of school) by parent reported questionnaire (validity r = .68, test-retest r=0.64) Total energy, fat, protein, carbohydrates intake by 3 day weighed food record-parent reported (r=NR)	Significant effect on leisure time MVPA and on total energy, fat, protein, saturated, monounsaturated and polyunsaturated fatty acids intake.	-Gender - Baseline values -SES -Initial BMI	NS NS NS NS	
Marcus[55] (2009) (low=4)	Sweden(STOPP) 10 schools n=1538 (PA), 770 (diet) %female= 50.5 Mean age=8.1 Drop-out=15.9% (PA), 10.1% (diet)	RCT I:4 school years, additional 30 min daily PA, changes at school lunch content, increasing awareness C: No intervention	PA (between 8 months-4 years post-baseline) Diet (4 years post-baseline)	-	Eating habits (FV, fast food, dairy products, nuts, sweets) by a questionnaire, parent-reported (r=NR) PA by 7 day uniaxis accelerometer (r=0.93, validity against direct calorimeter r=0.78-0.80)	Significant effect on high-fat dairy product, sweetened cereals, sweet products	- SES (low vs. high)	Significant on parent reported intake of dairy product and fast food	<u>Stratified by SES;</u> Low SES > High SES
McKenzie[57] (2001) (medium=5)	USA (CATCH) n= 96 elementary schools (School is the analysis unit) Drop-out= None	RCT I:2.5 year, PE modifications (curricular materials, development sessions for teachers) C: No intervention	Baseline; 6, 12, 18, 24, 30 and 36 months post-baseline	SCT	Physical activity in PE lessons by Systematic observation of PE lessons (r=0.94-0.99)	Significant effect on student's activity levels during PE lessons	-Teacher speciality (PE teacher vs. classroom teacher) -Location (indoor vs. outdoor)	Significant Significant	<u>Stratified by speciality;</u> Classroom teachers > PE teachers <u>Stratified by location;</u> Indoor > Outdoor for VPA, Outdoor> Indoor for walking
Perry[60] (1998) (medium=7)	USA (CATCH) 96 elementary schools, n= 1874 %female=50.0 Mean age=8.8 Drop-out=36.7%	RCT I:3 years, classroom curricula, food service changes, PE modifications C: No intervention	Baseline; 3 years post-baseline	SCT	FV consumption by 24-h recall (r=0.45-0.79)	NS	- Gender - Ethnicity - Site	NS NS NS	
Robinson[63] (1999) (medium=5)	USA 2 elementary schools n=198 %female=46.5 Mean age=8.9 Drop-out=3.0%	RCT I: 6 months, classroom curriculum, newsletters to parents C: No intervention	Baseline; 6 months post-baseline	SCT	Media use by questionnaire (r=0.94), Dietary intake (high-fat food) by 1-day food frequency recall (r=NR) Out-of-school PA by a checklist (r=0.76)	Significant intervention effect children's TV viewing, video game use and eating in front of TV	- Gender - Age	NS NS	

Shilts[65] (2009) (medium=5)	USA 1 middle school n=136 % female=45.0 Mean age=14.0 Drop-out=30.9%	RCT I: 5 weeks, guided goal setting (school curriculum, web-based assessment, workbook with handouts) C: Same intervention without goal setting	1 week before and 1 week after the intervention	SCT	Dietary and PA behavior by Centers for Disease Control Youth Risk Behavior Survey (YRBS) (dietary r=0.73, PA r=0.55)	ITT analysis: No significant effect Non-ITT analysis: Significant effects on dietary behavior scores	- Compliance (to the goal setting)	Significant for dietary behavior and PA	Stratified by compliance; High compliance > Low compliance
Singh[67] (2009) (medium=5)	The Netherlands (DoIT) 18 secondary schools n=1108 %female=53.3 Mean age=12.7 Drop-out=17.6%	RCT I: 8 months, Environmental component and Individual component (education program) C: No intervention	Baseline; 8, 12 and 20 months post-baseline	EnRG DPT TPB HT	Soft drinks and fruit juices consumption High-energy snack consumption, Screen-viewing Active transport	Significant effect on consumption of sugar containing beverages	- Gender - Ethnicity	Significant on screen viewing and soft drink consumption NS	Stratified by gender Screen viewing; Boys > Girls Soft drink consumption; Boys and girls both significant
Stevens[68] (2003) (medium=5)	USA (Pathways) 41 schools n=1455 %female=48.3 Mean age=7.5 Drop-out=29.5%	RCT I:3 year trial with a school curriculum, a PA/PE component, a school food service component, and a family component C: No intervention	Baseline; 3 years post-baseline	SLT	PA by Knowledge, attitude and behavior (KAB) questionnaire (r=NR)	Significant effect on physical activity	- Gender - Baseline weight status	NS NS	
Stock[69] (2007) (low=3)	Canada (Healthy Buddies) 2 elementary schools n=360 %female=55.3 Age range=5-12 Drop-out=35.8%	Quasi-experimental I;10 months, peer education by buddy lessons (presentations, games, art activities), PA sessions for buddy pairs (2/ per week) C: No intervention	Baseline; 10 months post-baseline	-	Healthy behavior score (includes healthy eating and PA) by Healthy Living Questionnaire (test-retest r=0.70-0.90)	Significant effect on health behavior among 4th through 7th grade	- Gender	Significant among students from kindergarten till 3rd grade	Stratified by gender Girls > Boys
Vandongen[73] (1995) (low=3)	Australia 30 schools n=1147 %female=51.3 Age range=10-12 Drop-out=24.2%	6-arm RCT I:9 months with 1- Physical fitness, 2-physical fitness+school nutrition,3-school nutrition,4-school + home nutrition, 5-home nutrition C:No intervention	Baseline; 9 months post-baseline	-	Total energy, fat, sugar, protein, fibre intake by 2 days dietary recall (r=NR)	NR	- Gender	Significant on Fat fibre and sugar intake	Stratified by gender; Fat and fiber intake: Girls > Boys Sugar intake: Boys>Girls

NS= Not significant , **NR**=Not reported, **FV**= Fruit and vegetable, **FJV**= Fruit, juice, vegetable, **SES**=Socio-economic status, **BMI**= Body mass index, **TV**=Television **PA**= physical activity, **LPA**= Light physical activity, **MVPA**=Moderate vigorous physical activity, **VPA**=vigorous physical activity, **PE**=Physical education, **SCT**=Social cognitive theory, **TM**= Transtheoretical model, **HPM**= Health promotion model, **SEF**=Social Ecological Framework, **TPB**=Theory of planned behavior, **ELM**=Elaboration likelihood Model, **SLT**=Social learning theory, **HBM**=Health Belief Model, **PBC**=Perceived behavioral control, **TRA**=Theory of reasoned action, **PBT**=Problem behavior theory, **HBPM**=Health Behavior Planning Model, **BCT**=Behavioral choice theory, **DPT**=Dual Process Theory, **HT**=Habit Theory, **EnRG**=EnRG framework, **ASE**=Attitude-Social Influence-Efficacy Model, **SM**=Self management, **TTI**= Theory of Triadic Influence, **OLT**= Operant learning theory, **OCT**=Organizational change theory, **DIM**= Diffusion of innovation model, **SAT**= Social Action Theory