Strength of obesity prevention interventions in early care and education settings: A systematic review

Dianne S. Ward a,⁎, Emily Welker b, Ashley Choate b, Kathryn E. Henderson c, Megan Lott b, Alison Tovar d, Amanda Wilson e, James F. Sallis e

a Department of Nutrition, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, United States
b Healthy Eating Research, Duke Global Health Institute, Duke University, Durham, NC, United States
c Henderson Consulting, Guilford, CT, United States
d Department of Nutrition and Food Sciences, University of Rhode Island, South Kingstown, RI, United States
e Department of Family Medicine and Public Health, University of California San Diego, San Diego, CA, United States

A R T I C L E   I N F O

Article history:
Received 3 May 2016
Received in revised form 20 September 2016
Accepted 23 September 2016
Available online 28 September 2016

Keywords:
Pediatric obesity
Child care
Child, preschool
Physical activity
Motor activity
Diet, food, and nutrition

A B S T R A C T

Time and place of study: 2010–2015; international. Given the high levels of obesity in young children, numbers of children in out-of-home care, and data suggesting a link between early care and education (ECE) participation and overweight/obesity, obesity prevention in ECE settings is critical. As the field has progressed, a number of interventions have been reviewed yet there is a need to summarize the data using more sophisticated analyses to answer questions on the effectiveness of interventions. We conducted a systematic review of obesity prevention interventions in center-based ECE settings published between 2010 and 2015. Our goal was to identify promising intervention characteristics associated with successful behavioral and anthropometric outcomes. A rigorous search strategy resulted in 43 interventions that met inclusion criteria. We developed a coding strategy to assess intervention strength, used a validated study quality assessment tool, and presented detailed descriptive information about interventions (e.g., target behaviors, intervention strategies, and mode of delivery). Intervention strength was positively correlated with reporting of positive anthropometric outcomes for physical activity, diet, and combined interventions, and parent engagement components increased the strength of these relationships. Study quality was modestly related to percent successful healthy eating outcomes. Relationships between intervention strength and behavioral outcomes demonstrated negative relationships for all behavioral outcomes. Specific components of intervention strength (number of intervention strategies, potential impact of strategies, frequency of use, and duration of intervention) were correlated with some of the anthropometric and parent engagement outcomes. The review provided tentative evidence that multi-component, multi-level ECE interventions with parental engagement are most likely to be effective with anthropometric outcomes.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Childhood obesity continues to be a global public health problem whereby the number of overweight or obese infants and young children (0 to 5 years) increased from 32 million globally in 1990 to 42 million in 2013 (Facts and Figures on Childhood Obesity, 2014). In the US, 22.8% of preschool aged children (2–5 years) were classified as overweight or obese (Ogden et al., 2014). Although obesity rates have recently decreased among this age group, racial/ethnic and socio-economic disparities continue (Ogden et al., 2014). The high rates and disparities are of concern, given that children who are overweight by age 5 are more likely to be obese later in life (Cunningham et al., 2014). To reduce lifetime risk of obesity, the Institute of Medicine recommends that obesity prevention interventions begin before the age of 5 (Early Childhood Obesity Prevention Policies, 2011).

Obesity-related diet and physical activity patterns of preschoolers do not meet national guidelines (Cortes et al., 2013; Hinkley et al., 2012; Kranz et al., 2006; Wilson et al., 2009). Children, especially those from racial/ethnic minorities and low-income communities in the U.S., eat too few fruits, vegetables, and whole grains, and consume too many energy dense snacks and beverages (Piernas and Popkin, 2011; Reedy and Krebs-Smith, 2010). Similarly, only half of preschool-aged children engage in the recommended 60 min of physical activity per day, and many exceed recommended limits for screen time, averaging 4 h per day (Beets et al., 2011; Tandon et al., 2011). Thus, interventions to improve eating and activity behaviors of preschool children are needed.

Although home environments are important for shaping eating and activity behaviors, >63% of U.S. mothers with preschool-aged children work outside the home (The State of America’s Children, 2013) and 70...
to 80% of children with working mothers spend on average 35 h per week in formal early care and education settings (ECE) (America’s Children: Key National Indicators of Well-Being, 2009; Larson et al., 2011a, 2011b; Ogden et al., 2012; Ward et al., 2008), mostly in center-based care (Child Care Costs on the Upswing, Census Bureau Reports, 2013). For children in full-time center-based care, approximately 50% of their daily dietary intake comes from meals and snacks served on site, and this location is the main source of their physical activity (Bollella et al., 1999; Gubbels et al., 2014; Padget and Briley, 2005). Given the numbers of children enrolled and the amount of time spent in this setting, promoting healthy eating and physical activity in ECE settings are integral to obesity prevention (Obesity in the Early Childhood Years, 2016). Because U.S. children who attend child care are at increased risk for obesity (Gubbels et al., 2010; Neelon et al., 2015; Woo Baidal et al., 2016), identifying successful interventions in these settings is critical, so they can be recommended for wide implementation.

Many reviews of obesity prevention efforts in ECE settings have been published (Blake-Lamb et al., 2016; Giampa et al., 2010; D’Onise et al., 2010; Hesketh and Campbell, 2010; Kreichauf et al., 2012; Larson et al., 2011a; Laws et al., 2014; Ling et al., 2016; Mikkelsen et al., 2014; Nelson et al., 2003; Nixon et al., 2012; Skouteris et al., 2011; Summerbell et al., 2012; Wolfenden et al., 2012), including a paper by Sisson et al. identified 71 ECE interventions, with more than two-thirds published since 2010. With this recent growth in ECE intervention studies, it is important for review papers to move beyond the single question of whether an intervention is generally effective and explore which specific characteristics and strategies contribute to intervention effectiveness.

In addition to adding a more comprehensive analysis, this review was also designed as a follow-up to the Larson et al. (2011a) review paper, which summarized child care-based intervention studies that covered the 10-year period between 2000 and 2010. The goal of the current study was to systematically review obesity prevention interventions in center-based ECE settings published 2010–2015 in order to identify the most promising intervention characteristics associated with successful behavioral and/or anthropometric outcomes. We hypothesized that more comprehensive and intensive interventions would be more effective. To accomplish this goal, we developed a coding strategy to assess intervention strength and allow for examination of several study questions:

1. Is intervention strength related to successful behavioral and/or anthropometric outcomes?
2. Are interventions that incorporate parent engagement more effective than those that do not?
3. Can specific intervention elements be identified that relate to desired outcomes, including number of intervention strategies used, potential impact of the strategies, and frequency and duration of these strategies?
4. Is overall study quality related to successful behavioral and/or anthropometric outcomes?

2. Methods

2.1. Search strategy

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were used (Moher et al., 2009). Separate searches (healthy eating, physical activity, and screen time) were conducted in three databases (PubMed, ERIC, and Web of Science) in December 2015, and search terms and methods are available in an online appendix (Appendix 2). Each search string contained four tailored components and corresponding terms: ECE setting; healthy eating, physical activity, or screen time; behavioral and anthropometric outcomes; and intervention-related. These searches returned 7494 results. After duplicates were removed, 6824 results were imported into EndNote for title and abstract review. Twenty-two additional papers were identified, nine of which were included, after cross-referencing included papers and recent reviews. Fig. 1 describes this process in more detail.

2.2. Inclusion and exclusion criteria

Papers were included if they were: peer-reviewed, published between 2010 and 2015, took place primarily in a center-based ECE setting, targeted children ages 0–6 years, included an intervention targeting healthy eating, physical activity and/or screen time, used an objective or validated measure of dietary intake, physical activity, screen-time, or anthropometric outcomes, provided a statistical measure of intervention success, and were published in English. All study designs, except case studies, were included if a pre- and post-evaluation was conducted.

2.3. Selection and data extraction

Two authors (EW & AC) each reviewed titles and abstracts, identifying 86 papers for full text review. EW and AC independently reviewed each article and extracted information using a template that included information on study design, location, sample characteristics, intervention components, and outcomes; meetings were held to discuss scoring and identify discrepancies. Any disagreements remaining following this discussion were resolved by consensus with a third reviewer who also read the full text (blinded to the other reviewers’ decisions). One author from a pool of five (ML, DW, KH, AW, and AT) reviewed summary entries to validate the extracted data and affirm inclusion of the paper. The goal of this review was to reach consensus; thus, statistics evaluating level of agreement were not computed.

2.4. Evaluation of intervention strength

For the purpose of this review we defined intervention strength as a composite of the number of intervention strategies used, their potential impact, and the frequency and duration of their use. This coding method is similar to ones used to characterize community health efforts for obesity prevention (Fawcett et al., 2015; Schwartz et al., 2015b). Individual intervention strategies were identified as described in the outcome paper and treated individually unless presented as part of a package (e.g., tool kit). Each intervention strategy was identified, evaluated for potential impact (scores 4 vs 1), and weighted by intensity (scores 1–4) and frequency (scores 1–4). Based on previous literature and ecological models (Sallis and Owen, 2015), high impact intervention strategies included changes in food, physical activity, and screen time environments or policies that provided ongoing support or structure for behavior change (e.g., revising menus or nutrition policies; providing more physical activity in the ECE setting; and in-person staff training). Low impact strategies included educational activities that encouraged individuals to make changes (e.g., field trips, information, posters, and games). High impact parent engagement included in-person strategies such as parent trainings or family days, while low impact strategies were more passive (e.g., sending materials home). Frequency of use was based on how often each strategy was employed. For example, if the intervention included the use of a weekly DVD, it was scored as a 3 on frequency (i.e., not daily but ≥1/week). A policy change such as increasing daily outdoor time by 30 min was scored as a 4 since it was in place during the entire intervention period. Duration of an intervention strategy was coded based on the length of the intervention period; thus there was one code per study. Two authors assessed intervention strength scores for each component (healthy eating, physical activity and screen time combined, and/or parent engagement). Disagreements were resolved by consensus. A cumulative strength score was then tallied. See Table 1 for a description of the coding scheme.
2.5. Assessment of intervention success

For each study, reported outcomes were coded in a systematic manner to determine intervention success. Dietary outcomes were not included if insufficient detail was provided on variety or quality of the food or food groups measured to determine desired direction (e.g., only grains consumed reported, but focus was increasing whole grains; only total milk consumption was reported, but without differentiation between full-fat and low/non-fat varieties). For physical activity outcomes, light physical activity was not coded. For anthropometric outcomes, only relative measures (e.g., BMI z-scores) were coded when both relative and absolute measures were reported. Both in-school and at-home behavioral outcomes were included. If studies included overlapping outcomes (e.g., MPA, VPA, and MVPA), all were coded. Sub-group analyses (e.g., boys, girls) were included when provided. When multiple analysis models were used, only outcomes from the adjusted models were included. To enhance comparability among studies, children were aware of research question was not coded.

1) overall intervention success (calculated as a percentage of successful outcomes to total number of outcomes ($\sum$ successful outcomes / $\sum$ all outcomes * 100)); and 2) any intervention success, where individual study success was a dichotomous outcome (any successful outcome vs. no successful outcome). These measures were calculated for healthy eating, physical activity/screen time, and anthropometric target areas if provided. The overall intervention success and dichotomous success scores (any) were used as the dependent variables in the hypothesis testing.

2.6. Quality assessment

Methodological quality of each study was assessed by two authors using the Quality Assessment Tool for Quantitative Studies from the Effective Public Health Practice Project (Quality Assessment Tool for Quantitative Studies, 2009). Ratings were compared and consensus reached on the final global rating. Given the nature of ECE studies where child care staff often provide intervention delivery, the question, “were study participants aware of research question?” was not coded.

<table>
<thead>
<tr>
<th>Number of strategies</th>
<th>Potential impact rating</th>
<th>Frequency rating</th>
<th>Duration rating</th>
<th>Strength score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tally of specific intervention strategies</td>
<td>4 = High (policy or environmental)</td>
<td>4 = Daily</td>
<td>4 = &gt; 9 months</td>
<td>$\sum$ of all intervention strategy score ratings</td>
</tr>
<tr>
<td>1 = Low (information, education)</td>
<td>3 = Not daily</td>
<td>3 = &gt; 6 months</td>
<td>Intervention strategy score rating = potential impact + frequency + duration</td>
<td></td>
</tr>
<tr>
<td></td>
<td>but $\geq$ 1 x/week</td>
<td>but $\leq$ 9 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 = $&lt; 1$ x/week</td>
<td>2 = $\geq$ 3 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>but $&gt; 1$ x/month</td>
<td>but $\leq$ 6 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 = $&lt; 3$ months</td>
<td>1 = $\leq 3$ months</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Study selection flow diagram. Time and Place of Study: 2010–2015; International. *Some of the papers included in the full text review were returned by two or all three of the search strings.
2.7. Analytic approach to hypothesis testing

All analyses were carried out in IBM SPSS Statistics for Windows, Version 23.0. Statistical tests included partial correlations between intervention strength scores and intervention success, adjusting for Quality Assessment score. Percent intervention success and intervention components (number of strategies, potential impact, frequency, and intervention duration) also were correlated. Study Quality Assessment scores and intervention success scores (continuous and dichotomous variables) were correlated as well. Finally, independent samples t-tests were conducted to compare studies with and without some key characteristics: child care staff vs externally-delivered. All statistical tests were two-tailed with alpha set to 0.05. However, these analyses are considered exploratory given the novelty of the review approach. Due to low sample sizes, correlations were interpreted when $r \geq 0.30$, regardless of p-value. Throughout the results, we note both statistical significance and this exploratory standard.

3. Results

3.1. Study characteristics

Papers ($n = 47$) included in this review described 43 unique interventions (see Table 2). Six additional papers describing intervention protocols were used to code intervention strength and quality. Of these interventions, over half took place in the US (33%), followed by Australia (14%), Germany (9%), and Switzerland (6%). The remaining studies took place in Switzerland, Chile, and Belgium (2 studies each), and England, Colombia, Spain, and Turkey (1 study each). Although some interventions were designed for any child enrolled in the center, measured child age in the studies ranged from 2+ to 6 years and were considered preschool interventions. Sample sizes ranged from 23 to 2062 children. The majority of papers included were RCTs (32), but other study designs included non-experimental pre-post (9), within and caregivers. Physical activity and screen-time outcomes were combined in the analyses.

3.2. Intervention characteristics

3.2.1. Intervention duration

Duration of the interventions ranged from 8 days to 3 years, with most lasting 4–6 months.

3.2.2. Mode of delivery

Approximately half of the interventions were delivered primarily by child care staff (56%). Other modes of delivery included: primarily by external experts (14%), primarily by the research team (12%), combination of child care staff and external experts (12%), or combinations of research staff, center staff, external experts, and parents (7%).

3.2.3. Target behaviors

Seven of the 43 interventions targeted healthy eating only, 17 targeted only physical activity, 9 targeted healthy eating and physical activity, 1 targeted physical activity and screen time, and 9 targeted healthy eating, physical activity, and screen time.

3.2.4. Type of strategies

Strategies to improve healthy eating included menu changes, nutrition education, changing meal service approaches, and food tastings. Strategies to improve physical activity included structured physical activity lessons, staff training, and take-home activity cards or resources. Twenty-five of the 43 interventions included a parental engagement component with strategies ranging from newsletters, CDs, or other handouts to more active strategies such as parent workshops or cooking classes. The median potential impact score sum for all strategies was 14 (range 4–39).

3.2.5. Number, frequency and duration of strategies used

The total number of strategies used within healthy eating, physical activity/screen time, and parental engagement ranged from 1 to 15, with 5 strategies as the median. The median strategies for healthy eating, physical activity, and parental engagement were 2 (range 0–6), 3 (range 0–9), and 1 (range 0–4) respectively. The average frequency of intervention strategies ranged from once a month or less to daily, with the median average frequency being less than once per week, but more than once per month.

3.2.6. Intervention strength score

The median intervention strength score was 40 (range 8–122). Strength scores for healthy eating and physical activity/screen time components were similar (median 23, range 5–51; and median 23.5, range 9–64, respectively). The median strength score for parent engagement was much lower at 15 (range 4–28).

3.3. Outcomes measured

Of the 43 interventions, 18 included one or more valid or objective measure for dietary intake, 31 for physical activity, 5 for screen time, and 24 for anthropometric outcomes. Measures of dietary intake included observation or plate waste (11), validated parent survey (5), food frequency questionnaires reported by parents (3), 24-h recalls (2), and food records (2). Measures of physical activity included accelerometers (16), motor development (10), direct observation (5), pedometers (3), obstacle course (1), and a questionnaire from the National Health and Nutrition Examination Survey (1). For anthropometric measures, the bulk of the studies used BMI (23), followed by body fat (7), waist circumference (4), waist-to-height ratio (1), weight (1), and mid-upper arm circumference (1). All five screen time measures were validated proxy reports from parents and caregivers. Physical activity and screen-time outcomes were combined in the analyses.

3.4. Study quality

Using the Quality Assessment Tool for Quantitative Studies (2009), 9 studies received a strong global rating, 14 received a moderate rating, and 20 received a weak rating.

3.5. Effect of interventions

3.5.1. Healthy eating

Of the 18 studies that included a dietary intake measure, the majority (72%) demonstrated at least one significant impact. Only five studies showed no effect on dietary intake. Percent successful outcomes ranged from 0 to 100%. Some studies demonstrated positive impacts on dietary intake, but only within specific food groups or nutrients such as fruits, vegetables, or sugar.

3.5.2. Physical activity/screen time

The majority (77%) of the 31 studies that measured changes in physical activity, fitness, or motor skills demonstrated at least one significant intervention effect. The remaining 7 studies showed no effect. Percent successful outcomes ranged from 0 to 100%, with a median of 50%. Only one of the 5 studies that measured screen time demonstrated a positive intervention effect.

3.5.3. Anthropometrics

Ten of the 24 studies with an anthropometric measure demonstrated at least one successful intervention effect. Percent successful outcomes ranged from 0 to 100%, with a median of 0%.
Table 2
Characteristics of included interventions, summary of results, quality assessment, and total intervention strength scores.

<table>
<thead>
<tr>
<th>First author and year</th>
<th>Study design</th>
<th>Location</th>
<th>Study design Duration</th>
<th>Characteristics of sample*</th>
<th>Exposure or intervention</th>
<th>Valid outcome type(s) and measures</th>
<th>Key/significant findings</th>
<th>Quality assessment/intervention strength score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alhassan et al. (2012)</td>
<td>Design: RCT</td>
<td>United States, MA</td>
<td>Duration: 6 months</td>
<td>Sample: 2 centers, 8 classrooms, N = 78 SES: Low R/E: 61% Latino/Hispanic, 39% AA Age: 2.9–5 yrs</td>
<td>LMS-based PA program including structured PA activities vs. unstructured free play time. Teacher trainings and lesson plans/resources provided.</td>
<td>Type: PA Measures: PA via accelerometer, and LMS via TGMD-2.</td>
<td>PA: Decrease in sedentary time and increase in leaping skills.</td>
<td>QA: Strong ISS: 26</td>
</tr>
<tr>
<td>Alhassan et al. (2013)</td>
<td>Design: RCT</td>
<td>United States, MA</td>
<td>Duration: 4 weeks</td>
<td>Sample: 2 centers, 8 classrooms, N = 75 Age: 2.9–5 yrs SES: Low R/E: NP</td>
<td>Based on SPARK, Normal outdoor play time + additional 30 min of structured outdoor play 3 days/week. Teacher trainings.</td>
<td>Type: PA Measures: PA via accelerometer.</td>
<td>PA: Decrease in sedentary time and increase in vigorous PA.</td>
<td>QA: Moderate ISS: 23</td>
</tr>
<tr>
<td>Alkon et al. (2014)</td>
<td>Design: RCT</td>
<td>United States, CA, CT, NC</td>
<td>Duration: 7 months</td>
<td>Sample: 17 centers, N = 552 Age: 3–5 yrs SES: Low R/E: 46% White, 17% Latino, 16% AA, 14% Asian, 7% other</td>
<td>Child Care Health Consultant worked with provider to write/update nutrition and PA policies. Provided NAP SACC workshops for center staff. On-site consultations, phone calls, emails, posters, info sheets.</td>
<td>Type: PA &amp; Anthro Measures: PA measured via OSRAP, BMI.</td>
<td>PA: NS. Anthro: Decrease in zBMI</td>
<td>QA: Weak ISS: 89</td>
</tr>
<tr>
<td>Annesi et al. (2013a)</td>
<td>Design: RCT</td>
<td>Southeast United States</td>
<td>Duration: 8 weeks</td>
<td>Sample: 32 classrooms, N = 275 Age: 3.5–5.6 yrs SES: Low-mid R/E: 100% AA</td>
<td>Start for Life Program. 30 min structured PA, 4 h of teacher training. Daily gross motor skills and behavior skill training, goal setting and self-monitoring with achievement charts, logs, and certificates. Activity binder. Start for Life Program</td>
<td>Type: PA Measures: PA measured via accelerometer.</td>
<td>PA: Improved MVPA and VPA in individuals and classrooms.</td>
<td>QA: Strong ISS: 30</td>
</tr>
<tr>
<td>Annesi et al. (2013b)</td>
<td>Design: RCT</td>
<td>Southeast United States</td>
<td>Duration: 9 months</td>
<td>Sample: 26 classrooms, N = 1154 Age: 4–5 yrs SES: Low - mid R/E: 86% AA, 9% Latino, 3% White, 2% other</td>
<td>Start for Life Program</td>
<td>Type: PA &amp; Anthro Measures: PA measured via accelerometer. BMI.</td>
<td>PA: Increased time in MVPA and VPA. Anthro: Decrease in BMI.</td>
<td>QA: Strong ISS: 40</td>
</tr>
<tr>
<td>Annesi et al. (2013c)</td>
<td>Design: RCT</td>
<td>Southeast United States</td>
<td>Duration: 9 months</td>
<td>Sample: 17 classrooms, N = 273 Age: 4–5 yrs SES: All at/below 130% of federal poverty line R/E: 100% AA</td>
<td>Start for Life Program</td>
<td>Type: PA &amp; Anthro Measures: PA measured via accelerometer. BMI.</td>
<td>PA: Greater % participating in MVPA and VPA. Anthro: Decrease in BMI.</td>
<td>QA: Strong ISS: 43</td>
</tr>
<tr>
<td>Baskale and Bahar (2011)</td>
<td>Design: Pre/post experimental with control</td>
<td>Location: Turkey</td>
<td>Duration: 1 yr</td>
<td>Sample: 6 schools, 2 from each income bracket, N = 238 Age: 5 yrs SES: Low, middle, high R/E: NP</td>
<td>Nutrition education intervention based on Piaget’s theory. Included play and visual materials, followed by observations of food selection and consumption.</td>
<td>Type: DI &amp; Anthro Measures: Food consumption frequency. MUAC. BMI.</td>
<td>DI: Improved white meat and fish, leafy and root vegetables, citrus fruits and other fruits, lower intake of sugar and processed fruit juice. Anthro: NS.</td>
<td>QA: Moderate ISS: 14</td>
</tr>
<tr>
<td>Bell et al. (2015)</td>
<td>Design: Pre/post cohort</td>
<td>Location: Australia</td>
<td>Duration: 1 month</td>
<td>Sample: 20 centers, N = 236 Age: 2–4 yrs SES: NP R/E: NP</td>
<td>The Start Right, Eat Right program. Teacher training and improved school healthy eating policies and menu.</td>
<td>Type: DI Measures: Plate waste</td>
<td>DI: Increased consumption of fruit, vegetables, protein, carbohydrates, minerals (Ca, Na, K, Mg, P, Zn), riboflavin, niacin, and folate. Decreased consumption of discretionary foods, fats and oils, % saturated fat.</td>
<td>QA: Weak ISS: 30</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>First author and year</th>
<th>Study design Location</th>
<th>Duration</th>
<th>Characteristics of sample</th>
<th>Exposure or intervention</th>
<th>Valid outcome type(s) and measures</th>
<th>Key/significant findings</th>
<th>Quality assessment/intervention strength score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonvin et al. (2013)</td>
<td>Design: Single-blinded RCT with 1:1 random assignment. Location: Switzerland</td>
<td>Duration: 9 months</td>
<td>Sample: 58 centers, N = 1467 Age: 3-5 yrs SES: Low parental EL R/E: 57.9% migrant</td>
<td>Youp'ta Bouge. Teacher training, rearrangement of the built environment, parental involvement, daily activity recommendation. Center received flyers, $1500, and advice. Centers encouraged to hold a parent information and discussion session. Parent flyers. Sesame Workshop Healthy Habits. Daily (1 h) classroom and play activities including storybooks, posters, videos, games, and songs. “Healthy family day” parent workshop, weekly newsletters. Teacher trainings, guide. Preventing Obesity by Design. Assistance to change the outdoor environment. POEMS site assessment. Seed grants. Teacher training workshops, webinars. TA. Website, renovation evaluation.</td>
<td>Type: PA &amp; Anthro Measures: PA measured using the Zurich Neuroromotor Assessment. PA measured using accelerometer in a subset of participants. BMI.</td>
<td>PA: Increase in MVPA, VPA and total PA in intervention group. Anthro: NS MVI: 36 mo: Decrease in overweight and increase in normal weight</td>
<td>QA: Moderate ISS: 41</td>
</tr>
<tr>
<td>Cespedes et al. (2013a, 2013b)</td>
<td>Design: RCT Location: Colombia</td>
<td>Duration: 5 months</td>
<td>Sample: 14 preschools, N = 1216 Age: 3-5 yrs SES: Low EL R/E: Low EL</td>
<td>Type: Anthro Measures: BMI.</td>
<td>Anthro: NS MVI: 36 mo: Decrease in overweight and increase in normal weight</td>
<td>QA: Moderate ISS: 66</td>
<td></td>
</tr>
<tr>
<td>Cosco et al. (2014)</td>
<td>Design: Pre/post Location: United States, NC</td>
<td>Duration: 1 month</td>
<td>Sample: 27 preschools, N = 804 Age: 3-5 yrs SES: Low EL R/E: Low EL</td>
<td>Type: PA Measures: PA measured via CARS</td>
<td>PA: Children were more likely to be engaged in non-sedentary activity. Connected single and double loops associated with higher activity than linear/straight pathways.</td>
<td>QA: Weak ISS: 15</td>
<td></td>
</tr>
</tbody>
</table>
### Table 2 (continued)

<table>
<thead>
<tr>
<th>First author and year</th>
<th>Study design</th>
<th>Location</th>
<th>Duration</th>
<th>Characteristics of sample</th>
<th>Exposure or intervention</th>
<th>Valid outcome type(s) and measures</th>
<th>Key/significant findings</th>
<th>Quality assessment/intervention strength score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hardy et al. (2010)</td>
<td>Design: Cluster RCT</td>
<td>Location: Australia</td>
<td>Duration: 20 weeks</td>
<td>Sample: 29 centers, N = 430</td>
<td>Age: 4.4 avg yrs SES: 47.7% Low R/E: 63% non-English spk</td>
<td>Munch and Move. Teacher training, resources, grant, and expert advice. Program included games, learning experiences, examples of policy statements, fact sheets etc.</td>
<td>Type: PA Measures: PA via the TGMD-2 checklist.</td>
<td>PA: Improved locomotor, object control and total FMS.</td>
</tr>
<tr>
<td>Harnack et al. (2012)</td>
<td>Design: Randomized crossover trial</td>
<td>Location: United States, MN</td>
<td>Duration: 6 weeks</td>
<td>Sample: 1 center, N = 57</td>
<td>Age: 2–5 yrs SES: 41.5% high school, 49.1% some college or assoc. degree R/E: 75.5% AA, 5.7% Hispanic/Latino, 13.2% multiracial, 3.8% American Indian, 1.0% White</td>
<td>Two serving styles: (1) Fruits and vegetables first: fruits and non-starchy vegetables served family style 5 min in advance of meal; and (2) Provider portioned meals: plate prepared for each child according to CACFP guidelines.</td>
<td>Type: DI Measures: Meal observation. Lunch observed data entered into NDSR.</td>
<td>DI: Fruit, vitamin A, and folate intakes were higher when fruits and vegetables served first. Fruit, vegetable, and total calorie intake higher in provider portioned meals.</td>
</tr>
<tr>
<td>Herman et al. (2012)</td>
<td>Design: Pre/post</td>
<td>Location: United States, PA, TX, AZ, RI, NY</td>
<td>Duration: 6 months</td>
<td>Sample: 75 centers, N = 112</td>
<td>Age: 3–5 yrs SES: Low parental EL R/E: Parents were 33.3% White, 32.4% Hispanic/Latino, 14.8% AA</td>
<td>Eat Healthy, Stay Active! Teacher training and TA provided. Parent training and incentives. Nutrition and PA lessons, and field trips to farmer’s market and grocery store for children.</td>
<td>Type: Anthro Measures: BMI</td>
<td>Anthro: Decrease in BMI and % obese.</td>
</tr>
<tr>
<td>Huss et al. (2013)</td>
<td>Design: Repeated exposure, randomized, cross-over quasi-experimental study</td>
<td>Location: United States</td>
<td>Duration: 12 weeks</td>
<td>Sample: 4 classrooms, N = 23</td>
<td>Age: 2–5 yrs SES: NP R/E: 56.5% Caucasian, 30.4% Asian, 13.1% other</td>
<td>Four combinations of portion size of main course and dessert: (1) reference portion, dessert with lunch; (2) reference portion, dessert after lunch; (3) large portion (50% larger), dessert with lunch; (4) large portion, dessert after lunch.</td>
<td>Type: DI Measures: Researchers measured plate waste and entered data into the NDSRParent survey.</td>
<td>DI: Serving dessert after meal increased energy intake from main course and dessert. Serving dessert with meal decreases total energy intake.</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>First author and year</th>
<th>Study design</th>
<th>Location</th>
<th>Duration</th>
<th>Characteristics of sample*</th>
<th>Exposure or intervention</th>
<th>Valid outcome type(s) and measures</th>
<th>Key/significant findings</th>
<th>Quality assessment/intervention strength score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monsalves Alvarez et al. (2015)</strong></td>
<td>Design: Cohort</td>
<td>Chile</td>
<td>Duration: 6 months</td>
<td>Sample: 1 school N = 70 Age: 3.2 avg yrs (boys); 3.3 avg yrs (girls) SES: NP R/E: NP</td>
<td>Physical activity classes including circuits with jumps, sprints, carrying medicinal balls, gallops, and crawling.</td>
<td>Type: PA &amp; Anthro Measures: Motor skill tests performed SLJ and 12 m run. BMI.</td>
<td>PA: Decrease in 12 m run time and increase in SLJ. Anthro: NS.</td>
<td>QA: Weak ISS: 17</td>
</tr>
<tr>
<td><strong>Natale et al., 2013, 2014</strong></td>
<td>Design: RCT</td>
<td>United States, FL</td>
<td>Duration: 6 months</td>
<td>Healthy Inside-Healthy Outside (HI-HO) program. Teacher training, weekly TA visits. Parent monthly dinner, monthly newsletters, and at-home activities. Schools developed new policies and menus to increase PA and healthy eating. Renovation of outdoor space. 1-shaped path transformed into a looping path, grassy hill was created, and climbing/sliding structures removed to create more open space.</td>
<td>Type: PA, DI, ST &amp; Anthro Measures: PA and ST questions were extracted from NHANES and modified (parent report); FFQ used for parents and teachers. BMI.</td>
<td>Type: PA, DI, ST &amp; Anthro Measures: PA measured via OSRAC-P and accelerometer.</td>
<td>DI: Decrease in junk food consumption, increase in mean fresh fruit and vegetable consumption. ST: Decrease in time spent at computer and TV. PA, NS, Anthro: NS.</td>
<td>QA: Strong ISS: 91</td>
</tr>
<tr>
<td><strong>Nicaise et al. (2012)</strong></td>
<td>Design: Cross sectional at two points</td>
<td>United States, CA</td>
<td>Duration: 1 yr</td>
<td>Sample: N = 50 pre, N = 57 post Age: 4–5 yrs SES: NP R/E: White (70.2% &amp; 63.2%), Hispanic (7.0% &amp; 13.2%), Asian (12.3% &amp; 13.2%), AA (10.4% &amp; 10.3%)</td>
<td>Teacher resource pack with 20 activity cards, user manual, lesson plans, sign posting information, and a poster promoting active play.</td>
<td>Type: PA Measures: PA via accelerometer.</td>
<td>PA: Increase in observed MVPA and decrease in observed sedentary time.</td>
<td>QA: Weak ISS: 12</td>
</tr>
<tr>
<td><strong>O'Dwyer et al. (2013)</strong></td>
<td>Design: Cluster RCT</td>
<td>United States</td>
<td>Duration: 6 weeks</td>
<td>Sample: 12 schools, N = 240 Age: 4.5 avg yrs SES: NP R/E: 84.3% White</td>
<td>Program S! Teacher training, classroom materials, online resources, and access to a blog to share best practices between schools. PA/DI lessons and take home activities. Annual health fair for families.</td>
<td>Type: PA, DI &amp; Anthro Measures: DI measured via questionnaire. BMI, skin fold, and waist circumference.</td>
<td>DI: NS. Anthro: NS.</td>
<td>QA: Moderate ISS: 89</td>
</tr>
<tr>
<td><strong>Penalvo et al. (2013a, 2013b, 2015)</strong></td>
<td>Design: Cluster RCT</td>
<td>Spain</td>
<td>Duration: 1–3 yrs</td>
<td>Sample: 24 schools, N = 2062 Age: 3–5 yrs SES: NP R/E: NP</td>
<td>Teacher training, classroom materials, online resources, and access to a blog to share best practices between schools. PA/DI lessons and take home activities. Annual health fair for families.</td>
<td>Type: PA, DI &amp; Anthro Measures: PA measured via accelerometer.</td>
<td>PA: Increase in activity, % MVPA &amp; % Total PA.</td>
<td>QA: Weak ISS: 20</td>
</tr>
<tr>
<td><strong>Roe et al. (2013)</strong></td>
<td>Design: Cross-over</td>
<td>England</td>
<td>Duration: 3–5 yrs</td>
<td>Sample: N = 61 Age: 3–5 yrs</td>
<td>Teachers offered children a healthy snack, either a</td>
<td>Type: DI Measures: Meal</td>
<td></td>
<td>QA: Weak ISS: 8</td>
</tr>
<tr>
<td>First author and year</td>
<td>Study design</td>
<td>Location</td>
<td>Duration</td>
<td>Characteristics of sample</td>
<td>Exposure or intervention</td>
<td>Valid outcome type(s) and measures</td>
<td>Key/significant findings</td>
<td>Quality assessment/intervention strength score</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------</td>
<td>----------</td>
<td>----------</td>
<td>---------------------------</td>
<td>-------------------------</td>
<td>---------------------------------</td>
<td>--------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Roth et al. (2010, 2015)</td>
<td>Design: Cluster RCT</td>
<td>Germany</td>
<td>1 yr</td>
<td>Sample: 41 schools, N = 709 Age: 4-5 yrs SES: Intervention group = 23.3% low, 50% mid, 26.7% high R/E: NP</td>
<td>Prevention through Activity in Kindergarten Trial (PAKT). Daily physical activity lessons, classroom materials, teacher training, and homework cards. Parents received interactive lectures and newsletters with health tips and activities.</td>
<td>Type: PA &amp; Anthro Measures: PA via accelerometer and obstacle course, BMI and skin fold measured.</td>
<td>PA: Increase in MVPA, motor skills performance, explosive leg strength, and jumping coordination. Anthro: Decrease in sum of four skin folds.</td>
<td>QA: Strong ISS: 56</td>
</tr>
<tr>
<td>Schwartz et al. (2015a)</td>
<td>Design: Within subjects crossover design</td>
<td>United States, CT</td>
<td>3 weeks</td>
<td>Sample: 1 center, N = 85 Age: 3-5 yrs SES: NP R/E: 81% Hispanic</td>
<td>Two variations of family-style feeding were compared to usual practice: (1) fruits, vegetables, and milk were served before the main meal (first course); and (2) fruits, vegetables, and milk were served before the main meal and meats and grains were removed from the table after the first serving (combination).</td>
<td>Type: PA &amp; DI Measures: PA via SOFIT-P. DI via meal observation.</td>
<td>PA: NS. DI: NS.</td>
<td>QA: Weak ISS: 31</td>
</tr>
<tr>
<td>Sharma et al. (2011)</td>
<td>Design: Convenience sample – pre/post pilot intervention</td>
<td>United States, TX</td>
<td>6 weeks</td>
<td>Sample: 2 centers, N = 75 Age: 3-5 yrs SES: Low R/E: Hispanic &amp; AA</td>
<td>Coordinated Approach to Child Health for Early Childhood (CATCH Early Childhood). There were four major components of the intervention: (1) Teacher-led, nutrition-based classroom curriculum “It’s Fun to Be Healthy!” (2) Teacher-led PA Box; (3) Parent education and tip-sheets; (4) Teacher training.</td>
<td>Type: PA &amp; DI Measures: PA via SOFIT-P. DI via meal observation.</td>
<td>PA: NS. DI: NS.</td>
<td>QA: Weak ISS: 16</td>
</tr>
<tr>
<td>Williams et al. (2014)</td>
<td>Design: Pre/Post RCT</td>
<td>United States, NY</td>
<td>6-10 weeks</td>
<td>Sample: 24 centers, N = 1143 Age: 3-5 yrs SES: NP R/E: 40% Latino, 24% White, 27% AA, 9% other</td>
<td>Eat Well Play Hard in Child Care. Part of NY SNAP-Ed. Multilevel messaging to preschool children, their parents, and center staff. RDNs selected 6 of 10 modules to use in classrooms and with parents separately. Healthy &amp; Ready to Learn. Child activities: set of children’s books and corresponding activities for parents/teachers to do with children. Increased daily MVPA. Teacher and parent training.</td>
<td>Type: PA Measures: TGMD-2 and Get Skilled, Get Active checklist. DI Measures: Parent survey of low-fat or fat-free milk; increase in vegetable intake. Increase in child-initiated vegetable snacking.</td>
<td>PA: Improved overall gross motor score and stationary ball kick. DI: More likely to drink low-fat or fat-free milk; increase in vegetable intake. Increase in child-initiated vegetable snacking.</td>
<td>QA: Strong ISS: 40</td>
</tr>
<tr>
<td>Winter and Sass (2011)</td>
<td>Design: Pre/Post quasi-exp matched sites</td>
<td>United States, TX</td>
<td>24 weeks</td>
<td>Sample: 4 centers, N = 405 Age: 3-5 yrs SES: Low R/E: 95% Latino</td>
<td>Color Me Healthy. Uses color, music, and exploration of the senses to teach children about healthy eating and PA. Circle</td>
<td>Type: DI Measures: DI measured via plate weight to calculate % of snack consumed.</td>
<td>DI: Increase in consumption of overall fruit, strawberry, cantaloupe, grape, pineapple, overall</td>
<td>QA: Weak ISS: 47</td>
</tr>
<tr>
<td>Witt and Dunn (2012)</td>
<td>Design: Randomized Pre/Post</td>
<td>United States, NY</td>
<td>3 weeks</td>
<td>Sample: 17 centers, N = 263 Age: 4-5 yrs SES: NP</td>
<td></td>
<td>Type: DI Measures: DI measured via plate weight to calculate % of snack consumed.</td>
<td>DI: Increase in consumption of overall fruit, strawberry, cantaloupe, grape, pineapple, overall</td>
<td>QA: Weak ISS: 41</td>
</tr>
</tbody>
</table>

(continued on next page)
Table 2 (continued)

<table>
<thead>
<tr>
<th>First author and year</th>
<th>Study design Location</th>
<th>Duration</th>
<th>Characteristics of sample</th>
<th>Exposure or intervention</th>
<th>Valid outcome type(s) and measures</th>
<th>Key/significant findings</th>
<th>Quality assessment/intervention strength score</th>
</tr>
</thead>
<tbody>
<tr>
<td>States, ID</td>
<td>Duration: 6 weeks</td>
<td>R/E: NP</td>
<td>time lessons and 1 imaginary trip each week. Take home activities for parents. Teacher training prior to program implementation on circle time lessons and imaginary trips.</td>
<td>vegetable, celery, cherry tomato, and broccoli. Fruit and veg snacks increased.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yin et al. (2012)</td>
<td>Design: Pre/Post quasi-experimental Location: United States, TX Duration: 18 weeks</td>
<td>Sample: 4 centers, N = 384 Age: 3–5 yrs SES: NP R/E: 90% Mexican-American</td>
<td>Míranos! Motor skill development, structured outdoor play, nutrition education and activities, integration of health literacy into classroom activities, staff development and wellness, and engagement of parents for support at home.</td>
<td>Type: PA, DI, ST &amp; Anthro Measures: PA via LAP-3 and pedometer on three consecutive days. DI via NDSR and aggregated plate waste measure. ST via parent survey. BMI.</td>
<td>PA: Improved gross motor development center and outdoor play intensity. DI: Improved consumption of fruits, vegetables, and low-fat milk. ST: NS. Anthro: Weight gain was less for combined center and home intervention.</td>
<td>QA: Moderate ISS: 88</td>
<td></td>
</tr>
</tbody>
</table>

AA: African American.
BMI: body mass index.
CACFP: Child and Adult Care Food Program.
CARS: Children’s Activity Rating Scale.
CMH: Color Me Healthy.
DI: dietary intake.
EL: education level.
FFQ: Food Frequency Questionnaire.
FMS: fundamental movement skills.
F/U: follow up.
HEL: high education level.
LAP-3: Learning Achievement Profile Version 3.
LEL: low education level.
LMS: locomotor skill.
M: migrant.
MoTB: Motor Test Battery.
MUAC: mid-upper arm circumference.
MVPA: moderate to vigorous physical activity.
N: sample size at baseline.
NAP SACC: Nutrition and Physical Activity Self-Assessment for Child Care.
NDSR: Nutrition Data System for Research.
NM: non-migrant.
NP: not provided.
NS: not significant.
OSRAC-P: Observational System for Recording Activity in Children-Preschool Version.
OSRAP: Observation System for Recording Activity in Preschools.
PA: physical activity.
POEMS: Preschool Outdoor Environment Measurement Scale.
PDMS: Peabody Developmental Motor Scales.
QA: quality assessment.
RDNs: registered dietitian nutritionists.
R/E: race/ethnicity.
SLJ: standing long jump.
SNAP: Supplemental Nutrition Assistance Program.
SORT-P: System for Observing Fitness Instruction Time for Preschoolers.
SPARK: Sports, Play, and Active Reaction for Kids.
ISS: Intervention Strength Score.
ST: screen time.
TA: teaching assistant.
TGMD: Test of Gross Motor Development.
VPA: vigorous physical activity.
WTHR: weight to height ratio.

---

* Sample size noted is based on the reported sample size at baseline; the final sample on which data analysis was performed may be smaller.

b Center no. based on Methods section; differed from abstract.
4. Hypotheses tested

4.1. Is intervention strength related to successful behavioral and/or anthropometric outcomes?

When correlating strength scores to percent successful outcomes/total outcomes, no significant positive correlation was observed for either a healthy eating or physical activity behavior outcome, with or without the inclusion of parent engagement (see Table 3). In fact, all of these correlations were unexpectedly in the negative direction. The strength of physical activity plus parent engagement and healthy eating plus parent engagement scores were both correlated greater than +0.30 with percent successful anthropometric outcomes, although this was not the case for the combined interventions.

When strength scores were correlated with the dichotomized outcome measure (some success vs. no success), physical activity intervention scores were correlated with physical activity outcomes > 0.30, but again, negatively. Correlations between healthy eating intervention strength scores and healthy eating outcomes were also unexpectedly negative, and the correlation between healthy eating plus parent engagement and any healthy eating outcome was greater than −0.30. By contrast, all six intervention strength scores were positively correlated with any successful anthropometric outcome, with correlations all > 0.30 and two reaching statistical significance (p < 0.05).

4.2. Are interventions that incorporate parent engagement more effective than those that do not?

Correlations of intervention strength scores with anthropometric outcomes were consistently higher when parent engagement scores were included. This pattern was stronger with the dichotomous anthropometric outcomes.

4.3. Can specific intervention elements be identified that relate to desired outcomes?

Of the 28 correlations shown in Table 4 relating number of strategies, potential impact, frequency, and duration to outcomes, seven were >0.30. Four of the seven were observed in parental engagement components.

4.4. Is the overall quality of the study related to desired behavioral and anthropometric outcomes?

As shown in the last row of Table 3, QA global rating score was positively correlated to both continuous and dichotomous outcomes for healthy eating, physical activity, and anthropometric outcomes, but only the healthy eating continuous correlation was >0.30.

Independent samples t-tests were conducted to compare RCTs with studies that did not meet RCT criteria on the primary outcomes (HE, PA, ST, anthropometric) for percent desired outcomes. Given the small sample sizes, none of the comparisons reached significance. However, we observed a consistent trend, with non-RCT studies showing a greater percent of positive outcomes across all four measures, with mean differences ranging from 7 to 16%.

We also compared those interventions delivered solely by childcare providers to those that were externally-delivered (i.e., external experts, research staff, or some combination). A trend was observed for externally-delivered interventions that showed a greater percent of positive outcomes across three of the four measures, with differences ranging from 9 to 26%. The anthropometric outcome, however, did not show this trend, with a difference of only 1.5% between the groups.

5. Discussion

Healthy eating, physical activity, and obesity prevention interventions in ECE settings have been an active area of study, with 43 interventions published between 2010 and 2015. Because of the extensive growth of interventions for obesity prevention (Larson et al., 2011b; Sisson et al., 2016), it is critical to understand which specific intervention characteristics and strategies contributed to intervention effectiveness. As such, we undertook a systematic review and created a coding scheme to examine the relation of intervention strength to healthy eating, physical activity/screen time, and anthropometric outcomes, based on the hypothesis that more intensive interventions should yield better outcomes.

In our review, we found that all of the intervention strength scores were correlated >0.30 with at least one significant anthropometric outcome. The strength scores included healthy eating interventions, physical activity/screen time interventions, and combined interventions. If this pattern is confirmed with further studies, the implication is that more comprehensive, multi-component, multi-level interventions with frequent and long-term implementation in ECE settings are more effective.
likely to be effective at helping children maintain or achieve healthy
weights. These findings are generally consistent with ecological models
of behavior (Sallis and Owen, 2015) and recommendations from the In-
ofstitute of Medicine (Accelerating Progress in Obesity Prevention:
Solving the Weight of the Nation, 2012) and other authoritative groups
favoring multi-level comprehensive interventions (Huang et al., 2009).

We found that parent engagement components added to the effec-
tiveness of ECE interventions. In all cases, correlations between inter-
vention strength and at least one significant anthropometric outcome
were higher when the strength of parent engagement was included in
the overall score (+0.16 for physical activity interventions; +0.15 for
health eating interventions; +0.10 for combined interventions, see
Table 3). Consistent with prior literature (Golley et al., 2011, Skouteris
health eating interventions; +0.10 for combined interventions, see
were higher when the strength of parent engagement was included in
Solving the Weight of the Nation, 2012) and other authoritative groups
stitute of Medicine (Accelerating Progress in Obesity Prevention:
vention strength and at least one signi
ficiency in the main findings with anthropometric outcomes
should be tempered by inconsistencies across the two methods of quan-
tifying outcomes. Prior reviews have categorized whether a study had
any significant intervention effect on a given outcome (Blake-Lamb et
al., 2016; Larson et al., 2011a; Laws et al., 2014; Ling et al., 2016;
Mikkelsen et al., 2014; Nixon et al., 2012; Sisson et al., 2016). Due to
concerns about overestimating success, we also computed the percent
of significant findings in the desired direction as a continuous measure
of outcomes. The latter approach was expected to roughly adjust for
the number of measures taken and tests conducted. The two approaches
produced mainly inconsistent results. Intervention intensity correla-
tions with the continuous outcome indicators were consistently lower
than correlations with the dichotomous anthropometric outcomes.
One possible explanation is that studies with multiple anthropometric
outcomes may have some measures with less sensitivity to change. Two
patterns seen with the dichotomous outcomes were generally rep-
licated with the continuous outcomes. Correlations with continuous an-
thropometric outcomes were higher when parent engagement was
included in the intervention strength score. Correlations with contin-
uous anthropometric outcomes tended to be higher for single-behavior
interventions compared to combined interventions. Only the physical
activity plus parent engagement and the healthy eating plus parent en-
gagement intervention scores were correlated with continuous anthrop-
ometric outcomes $>0.30$.

Our behavioral outcome findings were anomalous. For healthy eat-
ing intervention strength and dietary intake outcomes, all four corre-
lations were negative. Similarly, all four correlations between physical
activity intervention strength and physical activity outcomes were neg-
atively correlated. Thus, none of the findings supported the hypothesis
that more comprehensive ECE interventions produce better behavioral
outcomes. This pattern suggests a paradox whereby the strength of in-
terventions was correlated with less favorable behavioral outcomes
but more favorable anthropometric outcomes. The explanation that
measuring dietary and physical activity behaviors in young children is
imprecise is insufficient given that measurement error should bias cor-
relations toward zero, not negative.

We offer two potential explanations for these unexpected findings
related to behavioral outcomes. Visual inspection of the data indicated
that outliers may play a role in the negative correlations, which is not
surprising given the small number of studies. For example, the
scatterplot of physical activity intervention strength with physical activ-
ity continuous outcomes (not shown) demonstrated that the interven-
tion with the highest intervention strength had no significant interven-
tion effects, which appeared to enhance a negative correlation.
A more troublesome explanation may be that “stronger” interventions,
with multiple components and complex environmental and policy
changes over extended durations, may be particularly difficult to imple-
ment. The challenges of implementation may be amplified when child
care staff are instructed to deliver the interventions. Thus, we

### Table 4

<table>
<thead>
<tr>
<th>Components of intervention strength</th>
<th>Physical activity/screen time outcomes</th>
<th>Healthy eating outcomes</th>
<th>Anthropometric outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA/ST Intervention Strength Score</td>
<td>( N = 31 )</td>
<td>( N = 22 )</td>
<td></td>
</tr>
<tr>
<td>Strategies (#)</td>
<td>(-0.185)</td>
<td></td>
<td>0.423</td>
</tr>
<tr>
<td></td>
<td>( p = 0.320 )</td>
<td></td>
<td>( p = 0.050 )</td>
</tr>
<tr>
<td>Potential Impact Score</td>
<td>(-0.249)</td>
<td></td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td>( p = 0.177 )</td>
<td></td>
<td>( p = 0.684 )</td>
</tr>
<tr>
<td>Frequency Score</td>
<td>(-0.230)</td>
<td></td>
<td>0.276</td>
</tr>
<tr>
<td></td>
<td>( p = 0.212 )</td>
<td></td>
<td>( p = 0.213 )</td>
</tr>
<tr>
<td>Intervention Duration Score</td>
<td>(-0.094)</td>
<td></td>
<td>0.259</td>
</tr>
<tr>
<td></td>
<td>( p = 0.617 )</td>
<td></td>
<td>( p = 0.245 )</td>
</tr>
<tr>
<td>HE Intervention Strength Score</td>
<td>( N = 18 )</td>
<td>( N = 15 )</td>
<td></td>
</tr>
<tr>
<td>Strategies (#)</td>
<td>( 0.196 )</td>
<td>( 0.353 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p = 0.435 )</td>
<td>( p = 0.197 )</td>
<td></td>
</tr>
<tr>
<td>Potential Impact Score</td>
<td>( 0.232 )</td>
<td>( 0.149 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p = 0.355 )</td>
<td>( p = 0.595 )</td>
<td></td>
</tr>
<tr>
<td>Frequency Score</td>
<td>( 0.024 )</td>
<td>( 0.140 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p = 0.524 )</td>
<td>( p = 0.618 )</td>
<td></td>
</tr>
<tr>
<td>Intervention Duration Score</td>
<td>(-0.161)</td>
<td>( 0.303 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p = 0.524 )</td>
<td>( p = 0.273 )</td>
<td></td>
</tr>
<tr>
<td>PE Strength Score$^a$</td>
<td>( N = 18 )</td>
<td>( N = 13 )</td>
<td>( N = 20 )</td>
</tr>
<tr>
<td>Strategies (#)</td>
<td>( 0.352 )</td>
<td>( 0.294 )</td>
<td>( 0.052 )</td>
</tr>
<tr>
<td></td>
<td>( p = 0.152 )</td>
<td>( p = 0.330 )</td>
<td>( p = 0.826 )</td>
</tr>
<tr>
<td>Potential Impact Score</td>
<td>( 0.467 )</td>
<td>( 0.534 )</td>
<td>( 0.171 )</td>
</tr>
<tr>
<td></td>
<td>( p = 0.051 )</td>
<td>( p = 0.060 )</td>
<td>( p = 0.470 )</td>
</tr>
<tr>
<td>Frequency Score</td>
<td>( 0.121 )</td>
<td>(-0.065)</td>
<td>(-0.067)</td>
</tr>
<tr>
<td></td>
<td>( p = 0.633 )</td>
<td>( p = 0.833 )</td>
<td>( p = 0.778 )</td>
</tr>
<tr>
<td>Intervention Duration Score</td>
<td>( 0.103 )</td>
<td>( -0.154 )</td>
<td>( 0.348 )</td>
</tr>
<tr>
<td></td>
<td>( p = 0.684 )</td>
<td>( p = 0.616 )</td>
<td>( p = 0.133 )</td>
</tr>
</tbody>
</table>

PA/ST—physical activity/screen time; HE—healthy eating; PE—parent engagement.

$^a$ Includes only studies with parent intervention components.
hypothesize that comprehensiveness may be negatively associated with feasibility and fidelity of implementation. Physical activity intervention strength was more negatively related to behavioral outcomes than healthy eating intervention strength. This could be due to many healthy eating intervention components being implemented by food service workers who are more likely to have relevant training and support than child care workers whose wages are low and are generally not trained in providing physical activity. The critical role of implementation of prevention interventions was demonstrated in a review of 500 studies (Durlak and DuPre, 2008). Effect sizes were two to three times higher in studies with good implementation of interventions. Thus, it will be important to document and analyze implementation in future studies.

There was limited evidence that specific intervention components were particularly important. The negative (though nonsignificant) correlations between physical activity intervention scores and physical activity outcomes were repeated. Seven of 28 correlations between intervention components and outcomes was >0.30, with four drawn from the parent engagement components. Further research is needed to identify whether some intervention components contribute more to outcomes than other components. Also, there were no significant differences by study design (RCT vs. non-RCT) or by method of implementation (externally-delivered vs. child care staff).

Demonstrating increased interest in early childhood as an important time for intervention, several reviews on obesity prevention interventions in this age group have been published in recent years (Blake-Lamb et al., 2016; Ciampa et al., 2010; D’Onise et al., 2010; Hesketh and Campbell, 2010; Kreichauf et al., 2012; Larson et al., 2011a; Laws et al., 2014; Ling et al., 2016; Mikkelsen et al., 2014; Nelson et al., 2003; Nixon et al., 2012; Sisson et al., 2016; Skouteris et al., 2011; Summerbell et al., 2012; Wolfenden et al., 2012). The present paper overlaps in some ways with these reviews; however, it makes several unique contributions. First, previous work generally focused either on infancy (birth-two years) or early childhood (two-six years); however, the present review theoretically included children from birth to age six. It is important to note that no studies included in this review reported measurements for children ages birth to two years. Second, though several previous studies focused exclusively on the ECE setting (Kreichauf et al., 2012; Larson et al., 2011a; Mikkelsen et al., 2014; Nixon et al., 2012; Sisson et al., 2016; Zhou et al., 2014), many other reviews have included multiple intervention sites such as home, primary care, and mixed-setting interventions (Blake-Lamb et al., 2016; Ciampa et al., 2010; D’Onise et al., 2010; Hesketh and Campbell, 2010; Laws et al., 2014; Ling et al., 2016; Skouteris et al., 2011; Wolfenden et al., 2012). Considering the significant time that millions of young children spend in child care in the U.S. and globally, it was justified to focus only on interventions in ECE settings for the present review. Third, the present paper only included studies with objective or validated outcome measures. Fourth, several prior studies also examined the potential effects of parental involvement in ECE settings (Golley et al., 2011; Ling et al., 2016; Sisson et al., 2016; Skouteris et al., 2011; Sussner et al.). To our knowledge, this review is the first to code and quantitatively assess the impact of parental engagement intervention components on dietary intake, physical activity/screen time, and anthropometric outcomes. Fifth, previous reviews estimated the impact of interventions on outcomes by any positive effect (Blake-Lamb et al., 2016; Larson et al., 2011a; Laws et al., 2014; Ling et al., 2016; Mikkelsen et al., 2014; Nixon et al., 2012; Sisson et al., 2016). The present review makes a contribution by also examining the overall intervention success (calculated as the percentage of successful outcomes to total number of outcomes). Finally, the present review contributed a novel intervention strength coding tool and methods for incorporating those coding results into quantitative analysis to answer key questions about intervention characteristics and effectiveness.

An important strength of this review was the focus on examining the role of intervention strength in study outcomes. A coding system was developed to help capture the comprehensive, multi-level, multi-component nature of interventions given that they are the most likely to be effective (Sallis and Owen, 2015; Accelerating Progress in Obesity Prevention: Solving the Weight of the Nation, 2012). The coding was designed to roughly quantify components of interventions expected to predict better outcomes. By summarizing intervention components on a common metric, we were able to conduct a quasi-quantitative analysis that we hoped would be more informative than a narrative review. Given the lack of details presented in the published papers, the intervention strength metric may not have captured all of the study information. For example, there was not enough information to code specific intervention strategies across studies, and implementation of strategies was rarely reported. The content validity of the intervention strength coding was supported by similar indices used to assess childhood obesity intervention strength in the Healthy Communities Study (Fawcett et al., 2015) and to quantify intervention dose in community interventions (Schwartz et al., 2015b).

Though the intervention intensity coding system generated a wide range of scores, there was limited evidence to support the main hypothesis that ECE intervention strength is related to outcomes of healthy eating, physical activity and obesity prevention. Though nonsignificant, the pattern of findings supported tentative conclusions that intervention strength was positively correlated ($r \geq 0.30$) with the dichotomous anthropometric outcomes, and stronger parent engagement interventions tended to improve outcomes as noted by two significant correlations. These positive findings are particularly notable because the outcome measures were objective or validated. Our findings provide encouragement that well-designed ECE interventions may be able to play an important role in obesity prevention. However, other analyses did not support the intervention intensity hypothesis, raising the possibility that the positive results are not replicable. The negative correlations of intervention intensity scores and behavioral outcomes raise many questions. It was also surprising that the strength of single behavior interventions in general had higher correlations with anthropometric outcomes than did the combined diet and physical activity interventions, which could be related to issues of study power. Although our assessment of intervention strength is innovative and could add to the research literature, questions remain about our scoring. We encourage other investigators to develop alternate coding systems for intervention strength and compare results to the present approach.

Additional strengths of the study were the careful methods of the systematic review, use of a rigorous quality assessment method, and coding of study characteristics, study quality, and intervention strength by two raters. Several aspects of the review limit conclusions that can be drawn. We were unable to include intervention results from unpublished studies, a problem which has been recently noted among pediatric obesity interventions registered in the Clinical Trials Registration Database (Cui et al., 2015). To decrease publication bias, future intervention studies should always register trials with powered primary outcomes prior to study implementation. Due to the small number of studies with long-term follow-up ($n = 10$) and the variation in the follow-up time period (3 months to 3 years), only outcomes at the end of the intervention were analyzed. Measuring dietary and physical activity behaviors in young children is very challenging, so measurement limitations could explain some of the inconsistent and unexpected findings. The small number of studies used in correlations resulted in low statistical power, and several moderate-strength correlations (e.g., $r > 0.40$) were nonsignificant. Meta-analytic methods were not used, though a meta-analysis could not have evaluated the strength of the overall intervention and its components as was done here.

6. Conclusion

Preventing obesity in young children could have lifelong benefits, and many recent studies have evaluated obesity prevention interventions in preschool children. The exploratory, but quantitative, approach
to the present review of ECE interventions revealed that stronger interventions, with parent engagement and environmental and policy components, tended to be positively related to anthropometric outcomes. Thus, the best evidence suggests that comprehensive, multi-level obesity prevention interventions in ECE can be recommended. The present review raised several questions that should be priorities for future research.

- Important unanswered questions about the extent and quality of intervention implementation should be addressed in future studies, including the role of intervention complexity.
- Anomalous findings regarding intervention strength and behavioral outcomes should be examined further.
- Feasibility and effectiveness of single-behavior versus combined physical activity and healthy eating interventions requires more focused study.
- The intervention strength scoring system developed is similar to other systems (Fawcett et al., 2015), and we welcome other investigators to use and improve it.
- Importantly, it may be more productive to evaluate improved implementation of already-effective interventions than to study novel combinations of intervention strategies.

Conflict of interest

The authors declare no conflict of interest.

Transparency document

The Transparency document associated with this article can be found in the online version.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jspmed.2016.09.033.

References

Accelerating Progress in Obesity Prevention: Solving the Weight of the Nation. Institute of Medicine, Washington, D.C. (Available from: http://www.nationalacademies.org/


Importantly, it may be more productive to evaluate improved implementation of already-effective interventions than to study novel combinations of intervention strategies.

The authors declare no conflict of interest.

The Transparency document associated with this article can be found in the online version.


