Published in final edited form as:
Nutr Res. 2006 June ; 26(6): 241–248.

Children's dietary reporting accuracy over multiple 24-hour recalls varies by body mass index category

Suzanne Domel Baxter[a,*], **Albert F. Smith**[b], **Michele D. Nichols**[c], **Caroline H. Guinn**[a], and **James W. Hardin**[d]

a *Department of Health Education, Promotion, and Behavior; Arnold School of Public Health, University of South Carolina, Columbia, SC 29210, USA*

b *Department of Psychology, Cleveland State University, Cleveland, OH 44115, USA*

c *Center for Research in Nutrition and Health Disparities, Arnold School of Public Health, University of South Carolina, Columbia, SC 29204, USA*

d *Department of Epidemiology and Biostatistics and Center for Health Services and Policy Research, Arnold School of Public Health, University of South Carolina, Columbia, SC 29208, USA*

Abstract

This secondary analysis investigated the influence of body mass index (BMI) category and sex on reporting accuracy during multiple 24-hour dietary recalls. On three occasions, each of 79 children (40 girls) was observed eating school meals and interviewed the next morning about the previous day's intake, with ≥ 25 days between any two consecutive occasions for a child. Using age/sex BMI percentiles, we categorized 48 children as *healthy weight* (≥ 5th percentile <85th), 14 as *at risk of overweight* (≥ 85th percentile <95th), and 17 as *overweight* (≥95th percentile). A repeated-measures analysis was conducted for each of five outcomes (number of items observed eaten, number of items reported eaten, omission rate, intrusion rate, total inaccuracy). For *items observed*, BMI category x trial was marginally significant (*P*=0.079); over trials, this outcome was stable for healthy-weight children, decreased and stabilized for at-risk-of-overweight children, and was stable and decreased for overweight children. This outcome was greatest for overweight children and least for healthy-weight children (*P*=0.015). For *items reported*, no significant effects were found. For *omission rate* (*P*=0.028) and *intrusion rate* (*P*=0.083), BMI category x trial was significant and marginally significant; over trials, both decreased for healthy-weight children, decreased and stabilized for at-risk-of-overweight children, and increased and stabilized for overweight children. *Total inaccuracy* decreased slightly over trials (*P*=0.076); this outcome was greater for boys than for girls (*P*=0.049). Results suggest that children's dietary reporting accuracy over multiple recalls varies by BMI category. Validation studies with adequate samples for each BMI category, sex, and race are needed.

Keywords

Observations; Dietary recalls; Children; Sex; Body mass index category; Reporting accuracy

* Corresponding author: 220 Stoneridge Drive, Suite 103, Columbia, SC 29210, USA. Tel.: 803 251 6367 (extension 12); fax: 803 251 7954. E-mail address: sbaxter@gwm.sc.edu (S.D. Baxter)..

1. Introduction

Children have provided multiple 24-hour dietary recalls (24hDRs) over a period of several weeks or months for a variety of studies including national surveys [1], studies of heart disease etiology [2,3], evaluations of the effectiveness of nutrition education interventions [4–8], and evaluations of the relative validity of food frequency questionnaires [9–16]. However, to our knowledge, the consistency of children's reporting accuracy over multiple 24hDRs has been investigated in only one dietary validation study [17], conducted by our group. One finding from that study was that total inaccuracy (a measure that cumulates errors in servings across all items for each child [18]) decreased significantly from the first to the third interview; in other words, children's reporting accuracy for items and amounts combined improved over trials. This suggests that, in studies in which multiple 24hDRs are collected from children, the accuracy of the 24hDRs from individual children varies, and may improve, over trials.

Results from several studies with adults indicate that underreporting of energy intake increases as body mass index (BMI) increases [19–21], and that this is especially true for women [22–26]. Although some researchers [27] contend that in children, too, underreporting is related to BMI, studies that have assessed various dietary reporting methods relative to the metabolism of doubly-labeled water have provided conflicting results. Some studies with elementary school children (ages six to 11 years) have failed to find a relationship of dietary reporting accuracy to BMI [28–31], whereas others have found such a relationship [32–34]. Results from several studies with children [32–34] are consistent with studies of adults that indicate that underreporting increases as BMI increases. However, in some of these studies of children, parents assisted children in providing dietary reports [30,31,33]; parents provided the dietary reports [34]; children, parents, and observers provided the dietary reports [29]; or the extent to which parents and staff helped children provide dietary reports was unclear [28,32]. Thus, the dietary reports were not strictly by elementary school *children*, so the relationship between children's BMI and children's reporting accuracy cannot be determined from these cited studies.

In this article, we report additional analyses of the data collected in our study of the consistency of children's reporting accuracy over multiple interviews [17]. Specifically, we investigated whether the accuracy of fourth-grade children's reports of school meals (breakfast, lunch) over multiple 24hDRs was systematically related to BMI category and sex. Although the sample was stratified by race and sex, and children's height and weight were measured, the study was not designed to investigate associations of reporting accuracy over multiple 24hDRs with BMI category or sex, and we did not analyze BMI category or sex associations previously. We defined accuracy as the extent to which the parts of the 24hDRs concerning school meals corresponded to what the child was observed to have eaten at school breakfast and school lunch on the day about which the child was asked to report. Accuracy was assessed in terms of foods instead of kilocalories or nutrients because children (and adults) report what they have eaten as foods, not kilocalories or nutrients [35]. In the analyses presented in this article, we focused on whether dietary reporting accuracy improved or deteriorated consistently over trials.

2. Methods and Materials

The appropriate Institutional Review Board approved the study. Written child assent and parental consent to participate were obtained prior to data collection.

2.1. Sample

The sample has been described in detail elsewhere [17]. Briefly, during the 1999–2000 school year, fourth graders from six schools in one school district in Georgia were invited to participate; approximately 70% agreed. On three occasions, each of 79 randomly selected

NIH-PA Author Manuscript

NIH-PA Author Manuscript

children (20 black girls, 20 white girls, 19 black boys, 20 white boys) was observed eating school breakfast and school lunch, and then interviewed the following morning about the previous day's intake. Mean time between trial 1 and trial 2 was 46 days (SD = 15; range 28 – 99 days), and mean time between trial 2 and trial 3 was 41 days (SD = 13; range 25 – 89 days). In our earlier article [17], we reported data from 104 children, of whom 79 were observed and interviewed three times, 13 were observed and interviewed twice, and 12 were observed and interviewed once. For this article, we analyzed only data from the 79 children who were observed and interviewed three times each.

2.2. Observations

The observation procedure has been described in detail elsewhere [17]. Briefly, one of several observers stood by the tables at which the children were eating, and simultaneously observed one to three children. Only children who obtained their breakfast and lunch from the school foodservice were observed. Research staff observed children's entire meal periods so that traded foods could be noted [36,37]. Although children knew in general when they were being observed, individual children did not know who would be interviewed later. Practice observations were conducted prior to data collection to help familiarize children with an observer's presence. Inter-observer reliability, assessed during observer training and regularly throughout data collection, was acceptable [17].

2.3. Interviews

The interview procedure has been described in detail elsewhere [17]. Briefly, research staff who conducted observations also conducted 24hDR interviews, but an interviewer never interviewed a child that she had observed on the previous day. Interviews followed a four-pass protocol patterned after the Nutrition Data System for Research (version 4.03, Nutrition Coordinating Center, University of Minnesota, Minneapolis, 2000); however, information was written on a paper form. Each interview was audiorecorded and subsequently transcribed. Quality control for interviews, assessed during interviewer training and throughout data collection, indicated that interviewers adequately adhered to the protocol [17].

2.4. Measurement of Weight and Height

At approximately the mid-point of data collection, research staff used standardized procedures [38,39] to measure children's weight and height after lunch in a private location at school. Children were asked to remove their shoes and heavy clothing (e.g., jackets). Weight was measured to the nearest 1/10 pound using digital scales that were calibrated daily. Height was measured to the nearest 1/8 inch using a portable stadiometer. Inter-measurer reliability was assessed daily for pairs of research staff by having both staff members measure each of a 10% random selection of children; intraclass correlation reliability exceeded 0.99 for both weight and height.

2.5. BMI Categories

Each child's age at the time of measurement was calculated by subtracting his or her date of birth from the date of measurement. The Centers for Disease Control (CDC) sex-specific BMI-for-age growth charts for ages two to 20 years were used to determine each child's age/sex BMI percentile [40]. The sample distribution of age/sex BMI percentiles was skewed towards the high end; the mean ± standard deviation was 70 ± 26 for the 79 children, and the median was 77. Using the CDC categories, we classified children ≥ 5^{th} and <85^{th} percentiles as *healthy weight*, children ≥ 85^{th} and <95^{th} percentiles as *at risk of overweight*, and children ≥ 95^{th} percentile as *overweight* [41]. No child's BMI was below the 5^{th} percentile for his or her sex and age. Table 1 shows the number of children in each BMI category by sex and race, and the mean (± standard deviation) BMI for each BMI category.

2.6. Outcome Measures

We have explained in detail elsewhere [17] how we assessed the accuracy of children's reports of school meals. Briefly, to be counted as a report about a school meal, the child had to 1) indicate that the meal was eaten at school, 2) report the meal time to within an hour of the observed meal time, and 3) name the meal appropriately. Each item observed and/or reported eaten was classified as a match (an item observed eaten and reported eaten), an omission (an item observed eaten but not reported eaten), or an intrusion (an item reported eaten but not observed eaten). To be classified as a match, an item had to be reported as eaten at the school meal during which it was observed eaten. Because children can report foods many ways, we classified items reported eaten as matches unless it was clear that the children's reports did not describe items observed eaten; thus, we may have overestimated the accuracy of children's reports [17]. To construct the five outcome measures, items were weighted by importance according to meal component, with combination entrée = 2, condiment = 0.33, and other components = 1. Amounts observed and/or reported eaten were recorded in servings.

For each of a child's three observation/interview trials, five outcome measures were calculated for breakfast and lunch combined: number of items observed eaten, number of items reported eaten, omission rate, intrusion rate, and total inaccuracy. The legend for Table 2 defines these outcome measures. Note that for omission rate, intrusion rate, and total inaccuracy, *higher* values indicate *worse* reporting accuracy

2.7. Analyses

For each outcome measure, a full general linear model repeated-measures analysis was conducted. In each analysis, trial (1, 2, 3) was the within-subject factor (i.e., the repeated measure), and BMI category (healthy weight, at risk of overweight, overweight) and sex were between-subjects factors. The distribution of children into BMI categories varied over races (Specifically, 30% of black girls and 5% of white girls were overweight, and 42% of black boys and 10% of white boys were overweight.) Therefore, we included race as a covariate in each analysis. In these analyses, what was of major interest was the trend (improvement or deterioration), if any, in reporting accuracy over three trials, not the absolute level of reporting accuracy.

3. Results

Table 2 shows least-squares means (± standard errors) from the general linear model repeated-measures analysis conducted for each of the five outcome measures. Significant P values and F values are provided in the Table 2 legend.

For the number of items *observed* eaten, the BMI category x trial interaction was marginally significant ($P = 0.079$): Over three trials, the mean number of items observed eaten was stable for healthy-weight children, decreased and then stabilized for at-risk-of-overweight children, and was stable and then decreased for overweight children. In addition, BMI category was significant for the tests of between-subjects effects ($P = 0.015$): The mean number of items observed eaten was greatest for overweight children and least for healthy-weight children.

For the number of items *reported* eaten, no significant effects were found in the tests of within-subjects contrasts (all P values > 0.135), or in the tests of between-subjects effects (all P values > 0.247).

For *omission rate*, there was a significant effect of trial ($P = 0.006$): Over three trials, mean omission rate decreased and then stabilized. There was also a significant BMI category x trial interaction ($P = 0.028$): Over trials, mean omission rate decreased for healthy-weight children,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

decreased and then stabilized for at-risk-of-overweight children, and increased and then stabilized for overweight children.

For *intrusion rate*, there was a significant effect of trial ($P = 0.025$); the pattern was identical to that for omission rate: Over three trials, mean intrusion rate decreased and then stabilized. In addition, there was a marginally significant BMI category x trial interaction ($P = 0.083$), the pattern of which was identical to that for omission rate: Over trials, mean intrusion rate decreased for healthy-weight children, decreased and then stabilized for at-risk-of-overweight children, and increased and then stabilized for overweight children.

For *total inaccuracy*, there was a marginally significant effect of trial ($P = 0.076$): Over three trials, mean total inaccuracy decreased. In addition, sex was significant for the test of between-subjects effects ($P = 0.049$): Mean total inaccuracy was greater for boys than for girls.

4. Discussion

The purpose of these additional analyses was to investigate whether the accuracy of fourth-grade children's reports of school meals (breakfast, lunch) over multiple 24hDRs was systematically related to BMI category and sex. We found the following results: First, the accuracy of children's 24hDRs varied by BMI category over multiple recalls. For trial 1 only, overweight children had the lowest mean omission rate and intrusion rate. However, over three trials, mean omission and intrusion rates decreased for healthy-weight children, decreased and then stabilized for at-risk-of-overweight children, and increased and then stabilized for overweight children. Thus, the picture of children's reporting accuracy changed over time with multiple interviews according to BMI category. Specifically, over three trials, reporting accuracy improved for healthy-weight children, improved and then stabilized for at-risk-of-overweight children, and decreased and then stabilized for overweight children. These results are especially pertinent to studies that rely on children to complete multiple 24hDRs, for example, to evaluate the effectiveness of nutrition education interventions [4–8] or the relative validity of food frequency questionnaires [9–16].

Second, for no outcome measure did we find a significant interaction of BMI category and sex, or of trial and sex. For total inaccuracy, which cumulates errors for both items and amounts in servings, we found better reporting accuracy by girls than by boys. Because differences by sex were not found in omission rates or intrusion rates, we attribute the sex difference in total inaccuracy to differences between boys and girls in amounts reported for matches, amounts not reported for omissions, and/or amounts reported for intrusions. Classifying each item as a match, an omission, or an intrusion is a critical precursor to comparing observed amounts to reported amounts because amounts cannot be reported correctly if items are not reported correctly [17]. We were unable to conduct linear contrasts on the five outcome measures for each of the six BMI category x sex groups due to the small numbers of children in the two higher BMI categories.

Third, the number of items *observed* eaten at school breakfast and school lunch was significantly associated with BMI category—it was greatest for overweight children, and least for healthy-weight children (with at-risk-of-overweight children in between). In contrast, there was no significant association of BMI category with the number of items *reported* eaten at school breakfast and school lunch. These findings are of particular interest given the current attention to the increased prevalence of overweight among youth [42].

These additional analyses have several limitations. Most important, the original study was not designed to investigate the accuracy of children's dietary reporting over trials by BMI category and sex. Second, the confounding of race and BMI category (as indicated in Table 1) meant that race could be treated only as a covariate. Third, the number of girls and/or boys in two of the three BMI categories was small. Fourth, the number of days between any two consecutive reports by a child was uncontrolled, ranging from 25 to 99 (although this variation was not associated with BMI category or with sex).

We see these limitations as being more than offset by several strengths. The first—critical to the issue of any investigation of *children's* dietary reporting accuracy and BMI category—is that 24hDRs were obtained from children *without* assistance from their parents. This is important because it allowed us to identify errors in *children's* reporting accuracy. When parents assist children, or provide dietary reports instead of children, the extent to which children's dietary reporting errors are related to their own characteristics (e.g., BMI category; sex) cannot be determined.

Another strength of our study is that observations of school breakfast and school lunch were used to validate these parts of children's 24hDRs. Mertz asserted that observation is the best method for validating dietary reports, and recommended that observations occur in a cafeteria-type setting that is familiar to subjects [43]. Observations of school meals provide an attractive opportunity to validate parts of children's dietary reports in a setting that is familiar to them because millions of children eat school meals in a group setting on a regular basis [44–47]. Observations of children eating meals in private homes is intrusive [48], obvious [45], unacceptable in some communities [49], and may cause substantial reactivity [50]. However, reactivity is less problematic when observations are conducted at school [45] where children are accustomed to being watched while eating [45,51] and where groups may be observed in a way that keeps individual children from determining who, specifically, is being observed [52]. Validation studies that rely on the metabolism of doubly-labeled water are limited because they can identify only bias in reporting of energy intake: If dietary intake is underreported, doubly-labeled water cannot distinguish between underreporting across the dietary spectrum and selective underreporting of particular foods that might amplify bias in assessments of specific nutrients [27]. In contrast, validation studies that use observation are able to clarify whether underreporting is due to inaccurate reporting of items or of amounts, or both, and therefore can go beyond what studies using the metabolism of doubly-labeled water are able to provide.

Another strength of our study is that quality control for each aspect of data collection (observations, interviews, and measurements of weight and height) was assessed throughout data collection, and was acceptable. Studies often utilize several people to conduct direct observations; however, very few published studies using direct observations by multiple observers indicate whether inter-observer reliability was assessed [53]. Although the 24hDR is the most commonly used method for dietary surveys in the United States [54], few published studies that use 24hDRs mention quality control for interviews [55]. Measurement of weight and height is often included in studies, but few published studies mention assessment of inter-measurer reliability for measurements of weight and height.

A final strength of our study is that both aspects of reporting error – intrusion rate and omission rate – were assessed and analyzed. Although intrusion rate and omission rate are related by definition, they are empirically independent [35,56], and characterize different aspects of reporting accuracy: Omission rate is an input-bound measure that describes the accuracy of the reported proportion of the to-be-reported information while intrusion rate is an output-bound measure that describes the accuracy of what was reported [57,58]. For our study, the omission rate is the percentage of items observed eaten that was not reported; the intrusion rate is the percentage of reported items that was not observed eaten. Most attention to date has focused on underreporting (of amounts or omitted items), but the possibility of systematic overreporting (of amounts or intruded items) cannot be ignored [27]. Results that are limited to ascertaining whether the number of items reported increased (for example, over trials, or when evaluating a revised interview protocol) can be misleading because they indicate nothing about the extent to which the additional items were correctly reported (i.e., actually consumed) or falsely reported (i.e., intruded) [59]. In our analyses, the number of items reported eaten was not significantly associated with BMI category or sex, nor was there a significant BMI category x trial interaction. In contrast, for both omission rate and intrusion rate, we found a BMI category x trial interaction.

In summary, results from these additional analyses suggest that children's dietary reporting accuracy over multiple recalls varies by BMI category. In studies in which multiple 24hDRs are obtained from children, it cannot be assumed that reporting accuracy is invariant over trials and independent of BMI category. Dietary validation studies with adequate sample sizes of children for each BMI category, sex, and race need to be designed and conducted to extend and better understand these findings.

Acknowledgements

Funding was from grant R01 HL 63189 from the National Heart, Lung, and Blood Institute of the National Institutes of Health, and from grant 43-3AEM-2-80101 from the United States Department of Agriculture (Economic Research Service, FANRP). The authors thank the children, faculty, and staff of Goshen, Hephzibah, McBean, Monte Sano, Rollins, and Southside Elementary Schools; the School Nutrition Program; and the Richmond County Board of Education in Georgia for allowing data to be collected. The authors would also like to acknowledge Julie A. Royer, MSPH, for her assistance with revising the manuscript and tables.

References

1. Dwyer J, Picciano MF, Raiten DJ. Future Directions for What We Eat in America--NHANES: The Integrated CSFII-NHANES. J Nutr 2003;133:576S–81S. [PubMed: 12566506]
2. Van Horn LV, Stumbo P, Moag-Stahlberg A, Obarzanek E, Hartmuller VW, Farris RP, Kimm SY, Frederick M, Snetselaar L, Liu K. The Dietary Intervention Study in Children (DISC): Dietary assessment methods for 8- to 10-year-olds. J Am Diet Assoc 1993;93:1396–403. [PubMed: 8245373]
3. Snetselaar L, Stumbo P, Chenard C, Ahrens L, Smith K, Zimmerman B. Adolescents eating diets rich in either lean beef or lean poultry and fish reduced fat and saturated fat intake and those eating beef maintained serum ferritin status. J Am Diet Assoc 2004;104:424–8. [PubMed: 14993866]
4. Baranowski T, Henske J, Simons-Morton B, Palmer J, Tiernan K, Hooks PC, Dunn JK. Dietary change for cardiovascular disease prevention among Black-American families. Health Educ Res 1990;5:433–43.
5. Gortmaker SL, Cheung LWY, Peterson KE, Chomitz G, Cradle JH, Dart H, Fox MK, Bullock RB, Sobol AM, Colditz G, Field AE, Laird N. Impact of a school-based interdisciplinary intervention on diet and physical activity among urban primary school children: Eat Well and Keep Moving. Arch Pediatr Adolesc Med 1999;153:975–83. [PubMed: 10482216]
6. Perry CL, Mullis RM, Maile MC. Modifying the eating behavior of young children. J Sch Health 1985;55:399–402. [PubMed: 3853032]
7. Baranowski T, Baranowski J, Cullen KW, Marsh T, Islam N, Zakeri I, Honess-Morreale L, de Moor C. Squire's Quest! Dietary outcome evaluation of a multimedia game. Am J Prev Med 2003;24:52–61. [PubMed: 12554024]
8. Cullen KW, Himes JH, Baranowski T, Pettit J, Stevens M, Slawson DL, Obarzanek E, Murtaugh M, Matheson D, Sun W, Rochon J. Validity and reliability of a behavior-based food coding system for measuring fruit, 100% fruit juice, vegetable, and sweetened beverage consumption: Results from the Girls Health Enrichment Multisite Studies. Prev Med 2004;38:S24–S33. [PubMed: 15072856]
9. Cullen KW, Baranowski T, Baranowski J, Hebert D, de Moor C. Pilot study of the validity and reliability of brief fruit, juice and vegetable screeners among inner city African-American boys and 17 to 20 year old adults. J Am Coll Nutr 1999;18:442–50. [PubMed: 10511326]
10. Field AE, Colditz GA, Fox MK, Byers T, Serdula M, Bosch RJ, Peterson KE. Comparison of 4 questionnaires for assessment of fruit and vegetable intake. Am J Public Health 1998;88:1216–8. [PubMed: 9702152]

11. Field AE, Peterson KE, Gortmaker SL, Cheung L, Rockett H, Fox MK, Colditz GA. Reproducibility and validity of a food frequency questionnaire among fourth to seventh grade inner-city school children: Implications of age and day-to-day variation in dietary intake. Public Health Nutrition 1999;2:293–300. [PubMed: 10512564]
12. Rockett HRH, Breitenbach M, Frazier AL, Witschi J, Wolf AM, Field AE, Colditz GA. Validation of a youth/adolescent food frequency questionnaire. Prev Med 1997;26:808–16. [PubMed: 9388792]
13. Slater B, Philippi ST, Fisberg RM, Latorre MRDO. Validation of a semi-quantitative adolescent food frequency questionnaire applied at a public school in Sao Paulo, Brazil. Eur J Clin Nutr 2003;57:629–35. [PubMed: 12771963]
14. Speck BJ, Bradley CB, Harrell JS, Belyea MJ. A food frequency questionnaire for youth: Psychometric analysis and summary of eating habits in adolescents. J Adolesc Health 2001;28:16–25. [PubMed: 11137901]
15. Yaroch AL, Resnicow K, Davis M, Davis A, Smith M, Khan LK. Development of a modified picture-sort food frequency questionnaire administered to low-income, overweight, African-American adolescent girls. J Am Diet Assoc 2000;100:1050–6. [PubMed: 11019353]
16. Jensen JK, Gustafson D, Boushey CJ, Auld G, Bock MA, Bruhn CM, Gabel K, Misner S, Novotny R, Peck L, Read M. Development of a food frequency questionnaire to estimate calcium intake of Asian, Hispanic, and White youth. J Am Diet Assoc 2004;104:762–9. [PubMed: 15127061]
17. Baxter SD, Thompson WO, Litaker MS, Frye FHA, Guinn CH. Low accuracy and low consistency of fourth-graders' school breakfast and school lunch recalls. J Am Diet Assoc 2002;102:386–95. [PubMed: 11905461]
18. Baxter SD, Thompson WO, Davis HC. Prompting methods affect the accuracy of children's school lunch recalls. J Am Diet Assoc 2000;100:911–8. [PubMed: 10955049]
19. Braam LA, Ocke MC, Bueno-de-Mesquita HB, Seidell JC. Determinants of obesity-related underreporting of energy intake. Am J Epidemiol 1998;147:1081–6. [PubMed: 9620052]
20. Johnson RK, Soultanakis RP, Matthews DE. Literacy and body fatness are associated with underreporting of energy intake in US low-income women using the multiple-pass 24-hour recall: A doubly labeled water study. J Am Diet Assoc 1998;98:1136–40. [PubMed: 9787719]
21. Voss S, Kroke A, Klipstein-Grobusch K, Boeing H. Is macronutrient composition of dietary intake data affected by underreporting? Results from the EPIC-Potsdam Study: European Prospective Investigation into Cancer and Nutrition. Eur J Clin Nutr 1998;52:119–26. [PubMed: 9505157]
22. Briefel RR, Sempos CT, McDowell MA, Chien S, Alaimo K. Dietary methods research in the third National Health and Nutrition Examination Survey: Underreporting of energy intake. Am J Clin Nutr 1997;65:1203S–9S. [PubMed: 9094923]
23. Heitmann BL, Lissner L. Dietary underreporting by obese individuals - is it specific or non-specific? Br Med J 1995;311:986–9. [PubMed: 7580640]
24. Johnson RK, Goran MI, Poehlman ET. Correlates of over- and underreporting of energy intake in healthy older men and women. Am J Clin Nutr 1994;59:1286–90. [PubMed: 8198052]
25. Klesges RC, Eck LH, Ray JW. Who underreports dietary intake in a dietary recall? Evidence from the Second National Health and Nutrition Examination Survey. J Consult Clin Psychol 1995;63:438–44. [PubMed: 7608356]
26. Pryer JA, Vrijheid M, Nichols R, Kiggins M, Elliot P. Who are the 'low energy reporters' in the Dietary and Nutritional Survey of British Adults? Int J Epidemiol 1997;26:146–54. [PubMed: 9126514]
27. Livingstone MBE, Robson PJ. Measurement of dietary intake in children. Proceed Nutr Society 2000;57:279–93.
28. Bandini LG, Cyr H, Must A, Dietz WH. Validity of reported energy intake in preadolescent girls. Am J Clin Nutr 1997;65:1138S–41S. [PubMed: 9094910]
29. Champagne CM, Baker NB, DeLany JP, Harsha DW, Bray GA. Assessment of energy intake underreporting by doubly labeled water and observations on reported nutrient intakes in children. J Am Diet Assoc 1998;98:426–430, 433. [PubMed: 9550166]
30. Kaskoun MC, Johnson RK, Goran MI. Comparison of energy intake by semiquantitative food frequency questionnaire with total energy expenditure by the doubly labeled water method in young children. Am J Clin Nutr 1994;60:43–7. [PubMed: 8017336]

31. O'Connor J, Ball EJ, Steinbeck KS, Davies PSW, Wishart C, Gaskin KJ, Baur LA. Comparison of total energy expenditure and energy intake in children aged 6-9 y. Am J Clin Nutr 2001;74:643–9. [PubMed: 11684533]
32. Perks SM, Roemmich JN, Sandow-Pajewski M, Clark PA, Thomas E, Weltman A, Patrie J, Rogol AD. Alterations in growth and body composition during puberty. IV. Energy intake estimated by the Youth-Adolescent Food-Frequency Questionnaire: Validation by the doubly labeled water method. Am J Clin Nutr 2000;72:1455–60. [PubMed: 11101471]
33. Fisher JO, Johnson RK, Lindquist C, Birch LL, Goran MI. Influence of body composition on the accuracy of reported energy intake in children. Obes Res 2000;8:597–603. [PubMed: 11156436]
34. Maffeis C, Schutz Y, Zaffanello M, Piccoli R, Pinelli L. Elevated energy expenditure and reduced energy intake in obese prepubertal children: Paradox of poor dietary reliability in obesity? J Pediatr 1994;124:348–54. [PubMed: 8120702]
35. Smith, AF. Cognitive Processes in Long-term Dietary Recall. Hyattsville, MD: National Center for Health Statistics, Vital and Health Statistics; 1991. Series 6, No. 4
36. Baxter SD, Thompson WO, Davis HC. Trading of food during school lunch by first- and fourth-grade children. Nutr Res 2001;21:499–503.
37. Domel SB, Baranowski T, Leonard SB, Davis H, Riley P, Baranowski J. Accuracy of fourth- and fifth-grade students' food records compared with school-lunch observations. Am J Clin Nutr 1994;59:218S–20S. [PubMed: 8279428]
38. Lohman, TG.; Roche, AF.; Martorell, R. Anthropometric Standardization Reference Manual. Champaign, IL: Human Kinetics Books; 1988. p. 1-8.
39. Maternal and Child Health Bureau. Accurately weighing and measuring. [(accessed September 20, 2002; August 24, 2005; May 12, 2006).]. Available at http://depts.washington.edu/growth/
40. Kuczmarski RJ, Ogden CL, Guo SS, Bogus CH. 2000 CDC growth charts for the United States: Methods and development. National Center for Health Statistics. 2002[Vital Health Stat 11(246)]
41. Centers for Disease Control and Prevention, US Department of Health and Human Services. BMI – Body Mass Index: About BMI for Children and Teens. [(accessed August 30, 2005 and May 15, 2006).]. Available at http://www.cdc.gov/nccdphp/dnpa/bmi/childrens_BMI/about_childrens_BMI.htm
42. National Center for Health Statistics. Prevalence of Overweight among Children and Adolescents: United States, 1999-2002. [(accessed May 15, 2006).]. Available at: http://www.cdc.gov/nchs/products/pubs/pubd/hestats/overwght99.htm
43. Mertz W. Food intake measurements: Is there a "gold standard"? J Am Diet Assoc 1992;92:1463–5. [PubMed: 1452958]
44. Frank GC. Taking a bite out of eating behavior: Food records and food recalls of children. J Sch Health 1991;61:198–200. [PubMed: 1943041]
45. Simons-Morton BG, Baranowski T. Observation in assessment of children's dietary practices. J Sch Health 1991;61:204–7. [PubMed: 1943043]
46. US Department of Agriculture, Food and Nutrition Service. School Breakfast Program Annual Summary. [(accessed October 12, 2005 and May 15, 2006).]. Available at http://www.fns.usda.gov/pd/sbsummar.htm
47. US Department of Agriculture, Food and Nutrition Service. National School Lunch Annual Summary. [(accessed October 12, 2005 and May 15, 2006).]. Available at http://www.fns.usda.gov/pd/slsummar.htm
48. Caballero B, Davis S, Davis CE, Ethelbah B, Evans M, Lohman T, Stephenson L, Story M, White J. A school-based program for the primary prevention of obesity in American Indian children. J Nutr Biochem 1998;9:535–43.
49. Weber JL, Lytle L, Gittelsohn J, Cunningham-Sabo L, Heller K, Anliker JA, Stevens J, Hurley J, Ring K. Validity of self-reported dietary intake at school meals by American Indian children: The Pathways Study. J Am Diet Assoc 2004;104:746–52. [PubMed: 15127059]
50. Smith AF. Concerning the suitability of recordkeeping for validating and generalizing about reports of health-related information. Review of General Psychology 1999;3:133–50.

51. Baxter SD, Thompson WO, Litaker MS, Guinn CH, Frye FHA, Baglio ML, Shaffer NM. Accuracy of fourth-graders' dietary recalls of school breakfast and school lunch validated with observations: In-person versus telephone interviews. J Nutr Educ Behav 2003;35:124–34. [PubMed: 12773283]
52. Baxter SD, Thompson WO, Smith AF, Litaker MS, Yin Z, Frye FHA, Guinn CH, Baglio ML, Shaffer NM. Reverse versus forward order reporting and the accuracy of fourth-graders' recalls of school breakfast and school lunch. Prev Med 2003;36:601–14. [PubMed: 12689806]
53. Baglio ML, Baxter SD, Guinn CH, Thompson WO, Shaffer NM, Frye FHA. Assessment of interobserver reliability in nutrition studies that use direct observation of school meals. J Am Diet Assoc 2004;104:1385–93. [PubMed: 15354155]
54. Buzzard, M. 24-hour dietary recall and food record methods. In: Willett, W., editor. Nutritional Epidemiology. New York, NY: Oxford University Press; 1998. p. 50-73.
55. Shaffer NM, Baxter SD, Thompson WO, Baglio ML, Guinn CH, Frye FHA. Quality control for interviews to obtain dietary recalls from children for research studies. J Am Diet Assoc 2004;104:1577–85. [PubMed: 15389417]
56. Smith AF, Jobe JB, Mingay DJ. Retrieval from memory of dietary information. Appl Cognit Psychol 1991;5:269–96.
57. Koriat A, Goldsmith M. Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. J Exp Psychol Gen 1994;123:297–315. [PubMed: 7931094]
58. Koriat A, Goldsmith M. Monitoring and control processes in the strategic regulation of memory accuracy. Psychol Rev 1996;103:490–517. [PubMed: 8759045]
59. Baxter SD, Smith AF, Guinn CH, Thompson WO, Litaker MS, Baglio ML, Shaffer NM, Frye FHA. Interview format influences the accuracy of children's dietary recalls validated with observations. Nutr Res 2003;23:1537–46. [PubMed: 16724161]

Table 1

Number of fourth-grade children in each body mass index (BMI) category by sex and race, and mean (± standard deviation) BMI for each BMI category[a]

BMI category	BMI	Race	Sex: Girls	Sex: Boys	Total
Healthy weight	17.29 ± 1.44	Black	10	8	18
		White	14	16	30
		Category total	24	24	48
At risk of overweight	21.76 ± 1.04	Black	4	3	7
		White	5	2	7
		Category total	9	5	14
Overweight	26.31 ± 4.37	Black	6	8	14
		White	1	2	3
		Category total	7	10	17
		Overall total	40	39	79

[a]The Centers for Disease Control (CDC) sex-specific BMI-for-age growth charts for ages two to 20 years were used to determine each child's age/sex BMI percentile [40]. CDC categories were used to classify children as *healthy weight* (≥ 5th and <85th percentiles), *at risk of overweight* (≥ 85th and <95th percentiles), or *overweight* (≥ 95th percentile) [41].

Table 2

Least-squares means (± standard errors) from the general linear model repeated-measures analysis conducted for each of the five outcome measures (calculated for each day of a child's three days of observation/interview combinations)

		Number of items *observed* eaten [a, b]	Number of items *reported* eaten [a,c]	Omission rate (in %) [a,d]	Intrusion rate (in %) [a,e]	Total inaccuracy (in servings) [a,f,g]
BMI Category [h] x Trial:		[i]		[j]	[k]	
Healthy weight	Trial 1	8.3 ± 0.3	6.8 ± 0.3	56% ± 3%	47% ± 3%	7.5 ± 0.4
	Trial 2	8.1 ± 0.3	7.0 ± 0.3	48% ± 4%	39% ± 4%	6.7 ± 0.5
	Trial 3	8.4 ± 0.2	7.2 ± 0.3	42% ± 3%	33% ± 3%	6.2 ± 0.4
At risk of overweight	Trial 1	10.0 ± 0.5	6.8 ± 0.6	64% ± 6%	49% ± 6%	9.4 ± 0.8
	Trial 2	8.5 ± 0.5	6.7 ± 0.6	54% ± 7%	40% ± 7%	7.1 ± 1.0
	Trial 3	8.6 ± 0.4	6.5 ± 0.6	57% ± 6%	43% ± 7%	7.6 ± 0.8
Overweight	Trial 1	9.6 ± 0.5	7.3 ± 0.6	47% ± 6%	30% ± 6%	6.9 ± 0.8
	Trial 2	9.5 ± 0.5	7.1 ± 0.6	54% ± 6%	36% ± 7%	7.4 ± 0.9
	Trial 3	8.7 ± 0.4	6.2 ± 0.6	57% ± 6%	37% ± 6%	6.3 ± 0.8
Trial:				[j]	[k]	[l]
Trial 1		9.3 ± 0.2	7.0 ± 0.3	56% ± 3%	42% ± 3%	7.9 ± 0.4
Trial 2		8.7 ± 0.3	6.9 ± 0.3	52% ± 3%	38% ± 3%	7.1 ± 0.5
Trial 3		8.6 ± 0.2	6.6 ± 0.3	52% ± 3%	38% ± 3%	6.7 ± 0.4
BMI Category:		[i]				
Healthy weight		8.2 ± 0.2	7.0 ± 0.2	49% ± 2%	40% ± 2%	6.8 ± 0.4
At risk of overweight		9.0 ± 0.3	6.7 ± 0.5	58% ± 4%	44% ± 4%	8.0 ± 0.7
Overweight		9.3 ± 0.3	6.9 ± 0.4	53% ± 4%	34% ± 4%	6.9 ± 0.6
Sex:						[l]
Girl		8.6 ± 0.2	6.6 ± 0.3	53% ± 3%	38% ± 3%	6.6 ± 0.4
Boy		9.0 ± 0.2	7.1 ± 0.3	54% ± 3%	41% ± 3%	7.9 ± 0.5

[a]A weight was assigned to each match (an item observed eaten and reported eaten), omission (an item observed eaten but not reported eaten), and intrusion (an item reported eaten but not observed eaten) according to importance by meal component with combination entrée (e.g., hamburger, sausage biscuit) = 2, condiment (e.g., ketchup) = 0.33, and other meal components (e.g., chicken, milk, green beans, roll, pear, cake) = 1. Thus, reporting errors for combination entrées counted more than reporting errors for condiments and remaining meal components.

[b]Number of items observed eaten = sum of weighted number of items observed eaten at school breakfast and school lunch, per trial per child.

[c]Number of items reported eaten = sum of weighted number of items reported eaten for school breakfast and school lunch, per trial per child. For the weighted number of items reported eaten, no significant effects were found for the tests of within-subjects contrasts (all P values > 0.135), or for the tests of between-subjects effects (all P values > 0.247).

[d]For school breakfast and school lunch for each trial for each child, omission rate = (sum of weighted omissions / [sum of weighted omissions + sum of weighted matches]) * 100. Omission rates may range from 0% (i.e., no omissions) to 100% (i.e., no matches: no items observed eaten were reported eaten).

[e]For school breakfast and school lunch for each trial for each child, intrusion rate = (sum of weighted intrusions / [sum of weighted intrusions + sum of weighted matches]) * 100. The intrusion rate was undefined if a child was observed to eat something but reported eating nothing. Defined values for intrusion rates may range from 0% (i.e., no intrusions) to 100% (i.e., all reported items were intrusions).

[f]For school breakfast and school lunch for each trial for each child, total inaccuracy = (absolute difference between amounts reported and observed for each match * weight) + (amount of each omission * weight) + (amount of each intrusion * weight) summed over all items at observed meals on an individual day for each child. A total inaccuracy score of zero indicates a perfect recall relative to observation.

[g]Amounts eaten were observed, reported, and scored in servings as none (0.0), taste (0.1), little bit (0.25), half (0.50), most (0.75), all (1.0), or the actual number of servings if >1 serving (e.g., 2).

NIH-PA Author Manuscript
NIH-PA Author Manuscript

[h] According to the CDC sex-specific BMI-for-age growth charts and categories [40,41], children with BMI-for-age ≥ 5^{th} percentile and <85^{th} percentile were classified as *healthy weight*, children with BMI-for-age ≥ 85^{th} percentile and <95^{th} percentile were classified as *at risk of overweight*, and children with BMI-for-age ≥ 95^{th} percentile were classified as *overweight*.

[i] For the weighted number of items observed eaten, there was a marginally significant BMI category x trial interaction ($P = 0.079$; $F_{2,72} = 2.6$). In addition, BMI category was significant for the tests of between-subjects effects ($P = 0.015$; $F_{2,72} = 4.5$).

[j] For omission rate, there was a significant effect of trial ($P = 0.006$; $F_{1,72} = 8.1$) and a significant BMI category x trial interaction ($P = 0.028$; $F_{2,72} = 3.8$).

[k] For intrusion rate, there was a significant effect of trial ($P = 0.025$; $F_{1,72} = 5.3$) and a marginally significant BMI category x trial interaction ($P = 0.083$; $F_{2,72} = 2.6$).

[l] For total inaccuracy, there was a marginally significant effect of trial ($P = 0.076$; $F_{1,72} = 3.2$). In addition, sex was significant for the tests of between-subjects effects ($P = 0.049$; $F_{1,72} = 4.0$).