Published in final edited form as:

J Am Diet Assoc. 2002 March ; 102(3): 386–395.

Low accuracy and low consistency of fourth-graders' school breakfast and school lunch recalls

SUZANNE DOMEL BAXTER, PhD, RD, FADA

is an associate professor of nutrition at the Medical College of Georgia, Georgia Prevention Institute, Augusta.

WILLIAM O. THOMPSON, PhD

is director and professor emeritus, Medical College of Georgia, Office of Biostatistics and Bioinformatics, Augusta.

MARK S. LITAKER, PhD

is an assistant professor, Medical College of Georgia, Office of Biostatistics and Bioinformatics, Augusta.

FRANCESCA H.A. FRYE, RD and CAROLINE H. GUINN, RD

are research dietitians at the Medical College of Georgia, Georgia Prevention Institute, Augusta.

Abstract

Objective To determine the accuracy and consistency of fourth-graders' school breakfast and school lunch recalls obtained during 24-hour recalls and compared with observed intake.

Design Children were interviewed using a multiple-pass protocol at school the morning after being observed eating school breakfast and school lunch.

Subjects 104 children stratified by ethnicity (African-American, white) and gender were randomly selected and interviewed up to 3 times each with 4 to 14 weeks between each interview.

Statistical analysis Match, omission, and intrusion rates to determine accuracy of reporting items; arithmetic and/or absolute differences to determine accuracy for reporting amounts; total inaccuracy to determine inaccuracy for reporting items and amounts combined; intraclass correlation coefficients (ICC) to determine consistency.

Results Means were 51% for omission rate, 39% for intrusion rate, and 7.1 servings for total inaccuracy. Total inaccuracy decreased significantly from the first to the third recall ($P=0.006$). The ICC was 0.29 for total inaccuracy and 0.15 for omission rate. For all meal components except bread/grain and beverage, there were more omissions than intrusions. Mean arithmetic and absolute differences per serving in amount reported for matches were -0.08 and 0.24, respectively. Mean amounts per serving of omissions and intrusions were 0.86 and 0.80, respectively.

Applications/conclusions The low accuracy and low consistency of children's recalls from this study raise concerns regarding the current uses of dietary recalls obtained from children. To improve the

---

Correspondence to: SUZANNE DOMEL BAXTER.

Address correspondence to Suzanne Domel Baxter, PhD, RD, FADA, Associate Professor of Nutrition, Medical College of Georgia, Georgia Prevention Institute HS-1640, Augusta, GA 30912-3710.

This research was supported by grant HL 63189 from the National Heart, Lung, and Blood Institute of the National Institutes of Health. S.D. Baxter was the Principal Investigator.

The authors express appreciation to the children, faculty, and staff of Goshen, Hephzibah, McBean, Monte Sano, Rollins, and Southside Elementary Schools, and to the Richmond County Board of Education in Georgia for allowing data to be collected. Appreciation is also expressed to Candace Kopec, PhD, for her help with conducting observations and interviews, and to Roy Frye for developing the Microsoft Access forms and reports for data entry and data management.

accuracy and consistency of children's dietary recalls, validation studies are needed to determine the best way(s) to interview children.

The 24-hour dietary recall is the most commonly used method for dietary surveys in the United States (1) and is often used to collect information from children (2). For example, children's 24-hour recalls were used in the Bogalusa Heart Study (3), Child and Adolescent Trial for Cardiovascular Health (4), 5-a-Day Power Plus (5), Continuing Survey of Food Intakes by Individuals (6), and School Nutrition Dietary Assessment Study (7).

Validity is the extent to which a method provides accurate information. *Reliability* is the extent to which the information varies when the same method is administered on different occasions. Research concerning the validity and reliability of children's dietary recalls must include knowledge about foods actually eaten (8,9) because comparing information obtained from 2 self-report methods depends on the child's memory without knowledge of foods actually eaten. Although parents may provide information regarding what their children have eaten, several studies emphasize that once children begin attending school, parent's reports cannot be taken as truth (2,10,11). A better comparison is between information from a self-report method and information from a method independent of the child's memory, such as observation (8,9).

A recent Medline search (12) yielded 8 validation studies contained in 7 publications (10, 13-18) but no reliability studies regarding children's dietary recalls provided without parental assistance. To our knowledge, 9 other studies concern the validity of children's dietary recalls provided without parental assistance (11,19-26). Some studies (19,27) obtained children's dietary recalls on multiple days, and others (28-31) included "practice" recalls that were not analyzed. For one study (22) titled *Reliability and Validity of the 24-hour Recall*, children's recalls were compared with observations of 3 camp meals; however, only 1 recall per child was obtained. Thus, reliability pertained to the variance accounted for between observed and recalled values, which is different than consistency or variability from one recall to another. To our knowledge, no validation studies have evaluated the consistency of children's dietary recalls provided on multiple days without parental assistance. (The term *consistency* is used rather than *reliability* because each recall regards a different dietary event whereas the statistical term *reliability* regards measuring the same event multiple times.)

Observations of children eating school meals provide an excellent opportunity to validate portions of children's dietary recalls (32,33). Observations in homes may be too intrusive (34), but children are accustomed to being watched while eating at school. Foods eaten at school are important because a significant percentage of children's total daily intake is consumed at school (34). Regulations stipulate that school breakfast and school lunch provide one-fourth and one-third, respectively, of the daily recommended levels for energy, protein, calcium, iron, and vitamins A and C (35,36). More than 95% of children are enrolled in school (37); on a typical school day nationwide, almost 7.4 million and 27 million children participate in the School Breakfast and National School Lunch Programs, respectively (35,36).

Children report what they eat as foods, but the accuracy of dietary self-reports compared with actual intake is typically assessed indirectly at the nutrient level (38,39). Accuracy assessed indirectly may appear high for some nutrients but not others because substitutions of certain items may be similar to items actually eaten in some nutrients but not others (38,39). Insight gained from direct comparisons of foods reported eaten to foods actually eaten may guide research to improve methods for assessing diet to produce more accurate self-reports and provide practical guidance for eating (40-42).

The purpose of this study was to determine the accuracy and consistency of fourth-graders' school breakfast and school lunch recalls obtained during 24-hour recalls by comparing recalls

with observations. *Accuracy* was the extent to which a child's recalls provided correct information compared with observations. *Consistency* was the extent to which a child's accuracy varied from one day to another.

METHODS AND ANALYSES

Sample

The institutional Human Assurance Committee approved the study. References regarding reliability studies (43,44) were consulted to determine an appropriate design and sample size. The intraclass correlation coefficient (ICC) estimator was based on moment estimators of the between- and within- variance components (45). The expected length of a 95% confidence interval (45) for assumed ICCs between 0.4 and 0.6 for designs of 2 to 5 recalls per child was calculated. In all cases, the design with the shortest confidence interval per recall had 3 recalls per child. A sample size of 80 was selected because a sample of more children was of marginal benefit to further reducing the expected confidence interval width. Although the sample would be stratified by ethnicity and gender, the study was not powered to detect ethnic/gender differences.

Children were recruited from all 22 fourth-grade classes at 6 schools in one district during the 1999-2000 school year. The schools were selected to obtain a final sample with equal numbers of African-American (AA), white (W), male (M), and female (F) children with high participation in school breakfast and school lunch. A total of 36%, 66%, 68%, 71%, 83%, and 91% of the children across all grades at the respective 6 schools were eligible to receive free or reduced-price school meals during the data collection period. Of the 523 fourth-grade children (21% AAM, 22% AAF, 27% WM, 30% WF) invited to participate, 73% overall (n=382; 22% AAM, 23% AAF, 25% WM, 30% WF) provided child assent and parental consent (46). Of those who agreed to participate, a sample of 104 children, stratified by ethnicity and gender, was randomly selected.

Observations

Only randomly selected children who participated in school breakfast and school lunch were observed because contents of meals brought from home and eaten at school can be difficult to identify while conducting unobtrusive observations (47). One of 3 trained dietitians used a recording form while observing 1 to 3 children simultaneously. Observations followed previously used procedures and covered the entire breakfast and lunch periods to account for trading of foods (48,49). An observer stood by tables where children sat; thus, children knew when they were being observed, but did not know who would be interviewed. Practice observations were conducted with each class before data collection to acquaint children with the presence of observers and therefore lessen reactivity during data collection (33).

Inter-observer reliability (IOR) was conducted for training before data collection, and twice monthly throughout data collection. Results from IOR on 18 children indicated 89% agreement across observers for food items in which the amounts observed eaten were within one-fourth serving. This percent agreement is considered satisfactorily high (47,50).

24-hour recall interviews

Children were interviewed individually by 1 of 3 trained dietitians the morning after school breakfast and school lunch were observed. A different dietitian conducted the interview than the one who observed the child the previous day. Interviews were conducted in a private location at school, audio-recorded, and transcribed. For each child, a different dietitian conducted each interview on a different weekday when possible. Beginning and ending interview times were recorded to determine length. Before data collection, the minimum

number of weeks between any 2 recalls provided by each child was set at 4 to reduce possible learning effects; however, a maximum number was not set.

Interviews followed a written multiple-pass protocol. The original multiple-pass protocol was developed by the United States Department of Agriculture (51) and included 3 passes (52); they recently revised it to include 5 passes (53). The multiple-pass protocol used in this study had 4 passes (Figure 1) and was patterned after the one used by the Nutrition Data System for Research (NDS-R, version 4.03, Nutrition Coordinating Center, University of Minnesota, Minneapolis, 2000). Instead of using the computerized NDS-R version during the interviews, information was written on an interview form.

Inter-interviewer reliability (IIR) was conducted for training before data collection, and monthly throughout data collection. One dietitian interviewed a child while the remaining dietitian(s) sat behind the child and completed an interview form according to what was heard. At the end of the interview, the listening dietitian(s) could ask the child questions. Results from IIR on 8 children indicated that only one additional question would have been asked if a different dietitian had conducted the interview. The study's principal investigator also randomly selected and reviewed 5% of each interviewer's transcripts and audiotapes to ensure adherence to the interview protocol across dietitians.

Analyses

To determine which meals in the children's 24-hour recalls referred specifically to school breakfast and school lunch, children had to indicate school as the location where the meal was eaten, refer to breakfast as school breakfast or breakfast, and refer to lunch as school lunch or lunch. These requirements were established after determining they occurred in most recalls. Furthermore, the reported mealtime had to be within 1 hour of the observed mealtime. Although some children had difficulty reporting observed meal times exactly, especially at breakfast, most children could report times within 1 hour.

Items observed and/or reported eaten at school breakfast and school lunch were grouped by meal component and statistical weights were assigned (Figure 2). Combination entrees were considered a single meal component and counted only once during analyses. Each item reported and/or observed eaten was classified as a match, omission, or intrusion (Figure 2). Because foods can be reported many ways, items were scored as matches unless it was clear that the child's recall did not describe an observed food. This broad interpretation maximized the scored correctness of the child's recall; thus, true correctness may be overestimated. Examples of items observed and reported that matched were all types of white milk (eg, skim, 1%, whole) and all types of pizza (eg, cheese, sausage, pepperoni). Vegetables (eg, green beans, green peas), milk flavors (eg, chocolate, strawberry, white), and juices (eg, orange, apple) that differed were not matched.

To analyze accuracy for reporting food items (irrespective of amounts), matches, omissions, and intrusions were tallied for each recall for each child, and corresponding rates were calculated (16) (Figure 2). Omission and intrusion rates were calculated separately because a previous study indicated they are empirically uncorrelated (16). The omission rate addresses reality because it concerns the proportion of items eaten but not reported; the intrusion rate addresses the recall because it concerns the proportion of the recall that consists of items that were not observed (38,39). A child's recall could have low accuracy due to high omission and low intrusion rates, low omission and high intrusion rates, or high omission and high intrusion rates.

Amounts eaten were observed, recorded, and scored in servings (Figure 2). To analyze accuracy for reporting amounts for matches, absolute and arithmetic differences (16) were calculated

between amounts observed and reported eaten to determine the extent to which children underreported or overreported amounts (Figure 2). Amounts per item were calculated for omissions and intrusions separately (Figure 2) to assess whether these errors in reporting involved small or large amounts of servings.

A single measure of total inaccuracy (Figure 2), which was developed and evaluated previously (26), was used to capture children's total inaccuracy for reporting items and amounts combined. This measure has the advantage of being based on all items and amounts in the observation and/or recall; however, it fails to indicate whether errors are due to omissions, intrusions, or incorrectly reported amounts.

To determine the consistency (expressed as ICCs) of the accuracy of children's recalls, omission rate, intrusion rate, and total inaccuracy were each to be fit to a multivariate normal distribution with an assumed compound symmetric variance-covariance matrix (ie, a mixed-model analysis of variance) using SAS PROC MIXED. Total inaccuracy was square root transformed prior to analysis to stabilize the within-child variances across the range of mean inaccuracy per child. Data management and statistical calculations were conducted using Microsoft Access 2000 (SRI, Microsoft Corp, Redmond, Wash, 1999), SPSS for Windows (version 7.5, SPSS, Chicago, Ill, 1996) and SAS (Release 8.00, TS Level 00MO, SAS Institute, Inc, Cary, NC, 1999).

RESULTS

One hundred four children (24 AAM, 27 AAF, 25 WM, 28 WF) were observed and interviewed once each, 92 (21 AAM, 23 AAF, 23 WM, 25 WF) of the 104 children were observed and interviewed twice each, and 79 (19 AAM, 20 AAF, 20 WM, 20 WF) of the 92 children were observed and interviewed 3 times each for a total of 275 recalls to compare with school breakfast and school lunch observations. The time interval between each recall for each child ranged from 25 to 99 days (mean = 44; median = 41). Interviews to obtain the 24-hour recalls ranged in length from 5 to 29 minutes (mean = 15; median = 14).

Mixed-model analyses of variance, when applied to the square root of total inaccuracy and omission rate, both failed to indicate significant differences by interviewer or weekday. There was a significant sequence (first, second, or third recall) effect using square root of total inaccuracy ($P=0.006$). Least squares means (LSM) for the square root transformed data were 2.69, 2.58, and 2.45 for the first, second, and third recalls, respectively. Corresponding LSM for untransformed data were 7.54, 7.17, and 6.43 servings, respectively; these were slightly higher than if the transformed data were squared. The ICC for the square root of total inaccuracy was 0.29 ($P<0.0003$). Table 1 provides a descriptive summary of raw (unadjusted) means, medians, and modes for total inaccuracy, square root of total inaccuracy, omission rate, and intrusion rate for all 275 recalls.

Figure 3, Sections A and B, shows the distributions of omission and intrusion rates, respectively, across all 275 recalls. Results from a mixed-model analysis of variance on omission rates indicated an ICC of 0.15 ($P<0.04$), meaning that variability within child (ie, from one recall to another for the same child) was much greater than variability between children (ie, from one child to another). A mixed-model analysis of variance was not attempted on intrusion rates because the underlying structure of the data with 0% for 33 of the 275 recalls (12%) did not fit the required normality assumption. To further examine errors for reporting items, omission and intrusion rates were tallied by deciles for the 79 children with 3 recalls each (Table 2).

Across all 275 observations/recalls, there were 2,292 items observed eaten and 1,779 items reported eaten with 35% matches, 41% omissions, and 24% intrusions. Table 3 provides the

distribution of matches, omissions, and intrusions by meal component. For all meal components except for bread/grain and beverages, there were more omissions than intrusions.

For matches, mean arithmetic difference per serving in amounts reported was -0.08 , indicating an overall slight tendency to underreport amounts of items actually eaten; mean absolute difference per serving was 0.24 . Mean amounts of omitted and intruded items per serving were 0.86 and 0.80 , respectively.

DISCUSSION

The importance of validation studies has been thoroughly discussed (54) and the need for validated methods for assessing children's dietary intake has been emphasized repeatedly (12,25,34,55-59). However, this is the first validation study to evaluate the accuracy and consistency of children's dietary recalls provided on multiple days without parental assistance. Results indicate that the accuracy of children's school breakfast and school lunch recalls obtained during 24-hour recalls was poor compared with observation; furthermore, accuracy was inconsistent from one recall to another for the same child.

Considering that omission and intrusion rates may both range from 0% to 100%, with 0% as perfect, what constitutes acceptable accuracy? Applying an arbitrary pass or fail criteria for accuracy which establishes omission and intrusion rates of $\leq 30\%$ as passing and $>30\%$ as failing, only 3 of the 79 children (4%) achieved passing accuracy on each of their 3 recalls (Table 2). With this criteria, a child's recall would have acceptable accuracy even if the child omitted (ie, failed to report) up to approximately one third of items actually eaten, and if up to approximately one third of items reported by the child were intruded (ie, falsely reported). Taken together, a recall with this level of error would have little utility.

To what extent can recall accuracy vary from one recall to another for a child and still be deemed acceptably consistent? Applying an arbitrary consistent or inconsistent criteria that establishes omission and intrusion rates within 30% across an individual child's 3 recalls as consistent and $>30\%$ as inconsistent, only 18 of the 79 children (23%) provided consistent recalls (Table 2). With this criteria, recall accuracy would be consistent for a child who omitted 25%, 10%, and 35% of items and intruded 5%, 35%, and 15% of items during each of 3 respective recalls. However, recall accuracy would also be consistent for a child who omitted 80%, 70%, and 50% of items and intruded 50%, 75%, and 60% of items during each of 3 respective recalls. Thus, although acceptable consistency is important, acceptable consistency without acceptable accuracy is useless because children who are consistently inaccurate provide very little information regarding what they actually ate.

Analyzing reporting of items separately from reporting of amounts provides insight about what contributes to inaccurate recall, which in turn provides insight about what improvements can be made. When items are omitted or intruded, amounts cannot be correct. In this study, children omitted more than half of the items they were observed eating at school breakfast and school lunch; furthermore, of what they did report eating at these two meals, almost 40% was not observed eaten. Thus, research to improve the accuracy of children's dietary recalls must first focus on how to improve children's ability to correctly report items eaten. Furthermore, because children report what they eat as foods, not nutrients, unless reporting of items improves, derived measures of intake (eg, energy, nutrients) from children's recalls are not based on accurate reports of actual foods consumed.

When items were correctly reported (ie, matched), children's reported amounts were fairly accurate (ie, one-tenth and one-fourth serving for arithmetic and absolute differences, respectively). This is similar to results from a previous study with fourth-graders (16). The average omission was more than four-fifths serving; thus, omissions were generally not items

for which children ate only small amounts. The average intrusion was four-fifths serving; thus, intrusions were items for which children falsely claimed to have eaten most of the serving.

The total inaccuracy measure for each recall, based on all items observed and/or recalled across breakfast and lunch, captured the dimension of the total error, in servings, of the dietary recall. Across multiple days, this measure had an ICC of 0.29, indicating that there was less variability of total error from child to child than within children. We expected mean total inaccuracy to decrease (and thus accuracy of recall to increase) from the first to the third recall because some children quickly learn tasks and therefore perform better at them. Examining the within-child data, this learning effect was confined to a subset of children, suggesting that similar learning across all children in the population was unlikely; however, our sample size was insufficient to explore this supposition.

There were more omissions than intrusions for 8 of 10 meal components; this suggests that children were not simply substituting intrusions for omissions most of the time. The lowest omission rate by meal component was for beverage, perhaps because milk was regularly included for school breakfast and school lunch. Omissions and intrusions were high among all other meal components, including entree and combination entree. Although each child was asked during the third interview pass whether anything was added to each item (Figure 1), the omission rate was 67% for condiments, perhaps because children also omitted many items to which condiments were added.

The omission and intrusion rates from this study were higher than those from a previous study (16). Although both studies included fourth-graders, recalls in the previous study only included school lunch (16); in this study, however, recalls covered 24 hours. Thus, the cognitive burden of recalling school breakfast and school lunch items in the context of a 24-hour recall seems to be greater than that of recalling lunch items as a single meal, and negatively impacts fourth-graders' ability to accurately recall items eaten. This is similar to a comment by Reynolds et al (17) that underreporting is less of a problem if recalls are restricted to one meal.

To what extent did the interview protocol used in this study contribute to the low accuracy and low consistency of the children's recalls? This question cannot be answered because this study used the same interview protocol to obtain each recall from each child on each occasion. However, results from 3 validation studies (11,24,26), each with a small number of children, indicated that the structure of the interview protocol seemed to influence children's accuracy regarding daycare snack or school lunch intake compared with observation.

Several studies have compared children's dietary recalls with school lunch observations (eg, 10,13,15,16,24-26); however, to our knowledge, only one other study (14) also included school breakfast observations. Although certain aspects of school meals (eg, written cycle menu) may enhance memory for the items eaten (8,12), other aspects of school meals (eg, menu changes) may impede memory for items eaten (8).

An important strength of this study is the use of observation (which did not rely on memory) to validate children's recalls. This strength is especially important in understanding the low accuracy and low consistency of recalls from most children studied. If the recalls had been validated against another self-report method (eg, food diaries), many may have been deemed statistical artifacts rather than actual recalls with low accuracy and low consistency. Dietary recalls are commonly used with children, but research regarding the accuracy and consistency of children's recalls compared with non-self-report methods is scarce. Although there is concern that observations cause reactivity (54), if reactivity were an issue, the expected levels of accuracy would be higher than our results. Considering the importance of obtaining children's dietary recalls, research is needed to develop improved methods to increase the accuracy and consistency of their recalls. Such research needs to rely on non-self-report

methods for validation. Conducting observations of children's consumption during part of their day in an environment natural to them, such as school, provides a gold standard for comparison to their recalls.

Another strength of this study is that accuracy was analyzed several complementary ways. These included accuracy for reporting food items, accuracy for arithmetic and absolute differences per serving in reporting amounts for matches, errors in amounts per serving of omissions and intrusions, and total inaccuracy accounting for errors.

There are several limitations of this study. First, children were recruited from only 6 schools, which were not selected randomly; however, the percentages of AAM, AAF, WM, and WF children who agreed to participate were similar to those of the total population of fourth-graders at the 6 schools, which suggests that our sample is representative of the population (46). Second, the study was not powered to detect differences between the ethnic/gender groups; instead, the sample was stratified by ethnicity and gender. Third, analyses were limited to the school breakfast and school lunch portion of the 24-hour recall because these were the only meals observed.

APPLICATIONS

■ Dietitians and nutrition practitioners need to be cautious about how information regarding 24-hour dietary recalls obtained from children is used. Children in this study reported less than half of the items they were observed eating, and almost 40% of what they reported eating was not eaten. Furthermore, a child's accuracy could be high for one recall but low for another. Other objective evidence (eg, anthropometric measures, laboratory values) may provide some, but not all, of the missing information.

■ Nutrition researchers need to conduct validation studies to determine the best way(s) to interview children to enhance recall accuracy. Specifically, direction is needed regarding how to: a) initially instruct children to report items eaten, b) prompt children to report forgotten foods, and c) prevent children from falsely reporting items.

■ Dietitians and nutrition practitioners need to rely on concrete evidence to determine whether changes in methods for obtaining dietary recalls from children actually improve their accuracy and consistency.

References

1. Buzzard, M. 24-hour dietary recall and food record methods. In: Willett, W., editor. *Nutritional Epidemiology*. 2nd edition. Oxford University Press; New York, NY: 1998. p. 52
2. Presser, S.; Blair, J.; Mack, K.; Ryan, C.; Van Dyne, MA. Final report on the University of Maryland-USDA Cooperative agreement to improve reporting for children in the Continuing Survey of Food Intakes by Individuals. Survey Research Center, University of Maryland; Aug. 1993 Unpublished report
3. Nicklas TA, Webber LS, Srinivasan SR, Berenson GS. Secular trends in dietary intakes and cardiovascular risk factors of 10-y-old children: the Bogalusa Heart Study. *Am J Clin Nutr* 1993;57(6):930-937.
4. Luepker RV, Perry CL, McKinlay SM, Nader PR, Parcel GS, Stone EF, Webber LS, Elder JP, Feldman HA, Johnson CC, Kelder SH, Wu M. Outcomes of a field trial to improve children's dietary patterns and physical activity: The Child and Adolescent Trial for Cardiovascular Health (CATCH). *JAMA* 1996;275:768-776. [PubMed: 8598593]
5. Perry CL, Bishop DB, Taylor G, Murray DM, Mays RW, Dudovitz BS, Smith M, Story M. Changing fruit and vegetable consumption among children: The 5-a-Day Power Plus Program in St. Paul, Minnesota. *Am J Public Health* 1998;88:603-609. [PubMed: 9551002]

6. U.S. Department of Agriculture, Food and Nutrition Service, Office of Analysis, Nutrition and Evaluation. Changes in Children's Diets: 1989-1991 to 1994-1996, CN-01-CD2. Gleason, Phil; Sutor, Carol, editors. Project Officer, Ed Herzog; Alexandria, VA: 2001.
7. Burghardt JA. School Nutrition Dietary Assessment Study: overview of the study design. *Am J Clin Nutr* 1995;61(supplement):182S-186S. [PubMed: 7832164]
8. Domel SB. Self-reports of diet: How children remember what they have eaten. *Am J Clin Nutr* 1997;65(supplement):1148S-1152S. [PubMed: 9094912]
9. Fries E, Green P, Bowen DJ. What did I eat yesterday? Determinants of accuracy in 24-hour food memories. *Appl Cogn Psychol* 1995;9:143-155.
10. Emmons L, Hayes M. Accuracy of 24-hr recalls of young children. *J Am Diet Assoc* 1973;62:409-415. [PubMed: 4691953]
11. Mack, KA.; Blair, J.; Presser, S. Measuring and improving data quality in children's reports of dietary intake. In: Warnecke, R., editor. Health Survey Research Methods Conference Proceedings. DHHS publication; no. (PHS) 96-1013; 1996. p. 51-55.
12. McPherson RS, Hoelscher DM, Alexander M, Scanlon KS, Serdula MK. Dietary assessment methods among school-aged children: validity and reliability. *Prev Med* 2000;31:S11-S33.
13. Lytle LA, Nichaman MZ, Obarzanek E, Glovsky E, Montgomery D, Nicklas T, Zive M, Feldman H. Validation of 24-hour recalls assisted by food records in third-grade children. *J Am Diet Assoc* 1993;93:1431-1436. [PubMed: 8245378]
14. Todd KS, Kretsch MJ. Accuracy of the self-reported dietary recall of new immigrant and refugee children. *Nutr Res* 1986;6:1031-1043.
15. Lytle LA, Murray DA, Perry CL, Eldridge AL. Validating fourth-grade students' self-report of dietary intake: Results from the 5 A Day Power Plus program. *J Am Diet Assoc* 1998;98:570-572. [PubMed: 9597031]
16. Baxter SD, Thompson WO, Davis HC, Johnson MH. Impact of gender, ethnicity, meal component, and time interval between eating and reporting on accuracy of fourth-graders' self-reports of school lunch. *J Am Diet Assoc* 1997;97:1293-1298. [PubMed: 9366868]
17. Reynolds LA, Johnson SB, Silverstein J. Assessing daily diabetes management by 24-hour recall interview: The validity of children's reports. *J Ped Psychol* 1990;15:493-509.
18. Samuelson G. An epidemiological study of child health and nutrition in a northern Swedish county. *Nutr Metabol* 1970;12:321-340.
19. Bransby ER, Daubney CG, King J. Comparison of results obtained by different methods of individual dietary survey. *British J Nutr* 1948;2:89-110.
20. Meredith A, Matthews A, Zickefoose M, Weagley E, Wayave M, Brown EG. How well do school children recall what they have eaten? *J Am Diet Assoc* 1951;27:749-751. [PubMed: 14880399]
21. Greger JL, Etnyre GM. Validity of 24-hour dietary recalls by adolescent females. *Am J Public Health* 1978;68(1):70-72. [PubMed: 623366]
22. Carter RI, Sharbaugh CO, Stappell CA. Reliability and validity of the 24-hour recall. *J Am Diet Assoc* 1981;79:542-547. [PubMed: 7288060]
23. Stunkard AJ, Waxman M. Accuracy of self-reports of food intake. *J Am Diet Assoc* 1981;79:547-551. [PubMed: 7026651]
24. Domel SB, Thompson WO, Baranowski T, Smith AF. How children remember what they have eaten. *J Am Diet Assoc* 1994;94:1267-1272. [PubMed: 7963170]
25. Crawford PB, Obarzanek E, Morrison J, Sabry ZI. Comparative advantage of 3-day food records over 24-hour recall and 5-day food frequency validated by observation of 9- and 10-year-old girls. *J Am Diet Assoc* 1994;94:626-630. [PubMed: 8195550]
26. Baxter SD, Thompson WO, Davis HC. Prompting methods affect the accuracy of children's school lunch recalls. *J Am Diet Assoc* 2000;100:911-918. [PubMed: 10955049]
27. Rockett HR, Breitenbach M, Frazier AL, Witschi J, Wolf AM, Field AE, Colditz GA. Validation of a youth/adolescent food frequency questionnaire. *Prev Med* 1997;26(6):808-816. [PubMed: 9388792]

28. Van Horn LV, Gernhofer N, Moag-Stahlberg A, Farris R, Hartmuller G, Lasser VI, Stumbo P, Craddock S, Ballew C. Dietary assessment in children using electronic methods: Telephones and tape recorders. *J Am Diet Assoc* 1990;90:412–416. [PubMed: 2307817]
29. Achterberg C, Pugh MA, Collins S, Getty VM, Shannon B. Feasibility of telephone interviews to collect dietary recall information from children. *J Can Diet Assoc* 1991;52:226–228.
30. Dixon LB, McKenzie J, Shannon BM, Mitchell DC, Smiciklas-Wright H, Tershakovec A. The effect of changes in dietary fat on the food group and nutrient intake of 4- to 10-year-old children. *Pediatrics* 1997;100:863–872. [PubMed: 9346988]
31. McKenzie J, Dixon LB, Smiciklas-Wright H, Mitchell D, Shannon B, Tershakovec A. Change in nutrient intakes, number of servings, and contributions of total fat from food groups in 4- to 10-year-old children enrolled in a nutrition education study. *J Am Diet Assoc* 1996;96:865–873. [PubMed: 8784330]
32. Frank GC. Taking a bite out of eating behavior: food records and food recalls of children. *J School Health* 1991;61:198–200. [PubMed: 1943041]
33. Simons-Morton BG, Baranowski T. Observation in assessment of children's dietary practices. *J Sch Health* 1991;61:204–207. [PubMed: 1943043]
34. Caballero B, Davis S, Davis CE, Ethelbah B, Evans M, Lohman T, Stephenson L, Story M, White J. Pathways: A school-based program for the primary prevention of obesity in American Indian children. *J Nutr Biochem* 1998;9:535–543.
35. School Breakfast Program-Frequently Asked Questions. Accessed April 4, 2001. Available at: <http://www.fns.usda.gov/cnd/Breakfast/AboutBFast/faqs.htm>
36. School Lunch Program-Frequently Asked Questions. Accessed April 4, 2001. Available at: <http://www.fns.usda.gov/cnd/Lunch/AboutLunch/faqs.htm>
37. Kennedy E. Healthy meals, healthy food choices, healthy children: USDA's Team Nutrition. *Prev Med* 1996;25:56–60. [PubMed: 8778767]
38. Smith, AF. Vital and Health Statistics. 4. National Center for Health Statistics; Hyattsville, MD: 1991. Cognitive Processes in Long-term Dietary Recall. Series 6
39. Smith AF, Jobe JB, Mingay DJ. Retrieval from memory of dietary information. *Appl Cogn Psychol* 1991;5:269–296.
40. Baranowski T, Sprague D, Baranowski JH, Harrison JA. Accuracy of maternal dietary recall for preschool children. *J Am Diet Assoc* 1991;91:669–674. [PubMed: 2040780]
41. Smith AF. Cognitive psychological issues of relevance to the validity of dietary reports. *Eur J Clin Nutr* 1993;(Suppl 2):S6–S18. [PubMed: 8262022]
42. Willett, W. *Nutritional Epidemiology*. 2nd edition. Oxford University Press; New York, NY: 1998. p. 442
43. Donner A, Elasziv M. Sample size requirements for reliability studies. *Stat Med* 1987;6:441–448. [PubMed: 3629046]
44. Dunn, G. *Design and Analysis of Reliability Studies: a Statistical Evaluation of Measurement Errors*. Oxford University Press; New York: 1989.
45. Neter, J.; Kutner, MH.; Nachtsheim, CJ.; Wasserman, W. *Applied Linear Statistical Models*. 4th edition. Irwin; Chicago: 1996.
46. Frye FHA, Baxter SD, Thompson WO, Guinn CH. Recruitment methods and impact of school, class, ethnicity, and gender on agreement to participate in a nutrition study for fourth-graders. *J School Health*. In press
47. Simons-Morton BG, Forthofer R, Huang IW, Baranowski T, Reed DB, Fleishman R. Reliability of direct observation of school children's consumption of bag lunches. *J Am Diet Assoc* 1992;92:219–221. [PubMed: 1737906]
48. 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(supplement):218S–220S. [PubMed: 8279428]
49. Baxter SD, Thompson WO, Davis HC. Trading of food during school lunch by first- and fourth-grade children. *Nutr Res* 2001;21:499–503.

50. Baranowski T, Dworkin R, Henske JC, Clearman DR, Dunn JK, Nader PR, Hooks PC. The accuracy of children's self-reports of diet: Family Health Project. *J Am Diet Assoc* 1986;86:1381–1385. [PubMed: 3760429]
51. Guenther PM, DeMaio TJ, Berlin M. The multiple-pass approach for the 24-h recall in the Continuing Survey of Food Intakes by Individuals, 1994-1996. *Am J Clin Nutr* 1997;65(4supplement):1316S.
52. Johnson RK, Driscoll P, Goran MI. Comparison of multiple-pass 24-hour recall estimates of energy intake with total energy expenditure determined by the doubly labeled water method in young children. *J Am Diet Assoc* 1996;96:1140–1144. [PubMed: 8906138]
53. Moshfegh A, Borud L, Perloff B, LaComb R. Improved method for the 24-hour dietary recall for the use in national surveys. *FASEB J* 1999;13(4):A603. Also available at <http://www.barc.usda.gov/bhnrc/foodsurvey/pdf/method.pdf> Accessed 6/20/2001
54. Smith AF. Concerning the suitability of recordkeeping for validating and generalizing about reports of health-related information. *Rev Gen Psychol* 1999;3(2):133–150.
55. Champagne CM, Delany JP, Harsha DW, Bray G. Underreporting of energy intake in biracial children is verified by doubly labeled water. *J Am Diet Assoc* 1996;96:707–709. [PubMed: 8675915]
56. Contento IR, Manning AD, Shannon B. Research perspective on school-based nutrition education. *J Nutr Educ* 1992;24:247–260.
57. Lytle, LA. Nutrition Education for School-aged Children: a Review of Research. US Dept of Agriculture, Food and Consumer Service, Office of Analysis and Evaluation; Washington, DC: 1995.
58. Randall E. Measuring food use in school-aged children. *J School Health* 1991;61:201–203. [PubMed: 1943042]
59. Van Horn LV, Stumbo P, Moag-Stahlberg A, Obarzanek E, Hartmuller VW, Farris RP, Kimm SYS, Frederick M, Snetselaar L, Lui K. The Dietary Intervention Study in Children (DISC): dietary assessment methods for 8- to 10-year-olds. *J Am Diet Assoc* 1993;93:1396–1403. [PubMed: 8245373]

FIG 1.
Description of multiple-pass protocol used for interviews.

FIG 2.**Definitions.**

^aA statistical weight was assigned to each item (match, omission, or intrusion) by meal component [entree, combination entree, bread/grain, vegetable, fruit, dessert, condiment, beverage, breakfast meat, and miscellaneous (foods such as candy and chips that did not fit into any of the other components)]. Statistical weights were combination entree (eg, taco salad, sausage biscuit) = 2, condiment (eg, ketchup) = 0.33, and remaining meal components = 1. Statistical weights were used so that errors regarding reporting combination entrees counted more than errors for condiments and remaining meal components.

FIG 3.
Distribution of omission rates (section A) and intrusion rates (section B) across all 275 recalls from 104 children.

Table 1

Means, medians, and modes for total inaccuracy, square root of total inaccuracy, omission rate, and intrusion rate for all 275 recalls from 104 children

Outcome variable	Mean	Median	Mode
Total inaccuracy ^{a,b} (in servings)	7.1	6.8	3.3
Square root of total inaccuracy ^{a,c}	2.6	2.6	1.8
Omission rate (%) ^{a,d}	51	53	50
Intrusion rate (%) ^{a,e}	39	40	0

^a A statistical weight was assigned to each item by meal component with combination entree=2, condiment=0.33, remaining meal components=1.

^b Total inaccuracy=(absolute difference between amounts reported and observed eaten for each match×statistical weight)+(each omitted amount×statistical weight)+(each intruded amount×statistical weight) summed over all items at school breakfast and school lunch on an individual day for each child. A score of zero indicated a perfect recall compared with observation.

^c Square root transformation of total inaccuracy.

^d Omission rate=[sum of weighted omissions/(sum of weighted omissions+sum of weighted matches)]×100. Omission rates were undefined if a child was observed to have eaten nothing regardless of what was reported eaten. Omission rates ranged from 0% to 100%, with 0% considered perfect.

^e Intrusion rate=[sum of weighted intrusions/(sum of weighted intrusions+sum of weighted matches)]×100. Intrusion rates were undefined if a child was observed to have eaten something yet failed to report eating anything. Intrusion rates ranged from 0% to 100%, with 0% considered perfect.

Table 2

Distribution of omission, intrusion, and omission and intrusion rates across 79 children with three recalls each

Deciles	Omission rate ^{a,b}		Intrusion rate ^{a,c}		Omission and intrusion rates ^{a,d}	
	within ^e	≤ ^f	within ^e	≤ ^f	within ^e	≤ ^f
	← cumulative no. children →					
10%	9	0	13	1	3	0
20%	30	1	26	5	5	1
30%	44	3	45	9	18	3
40%	58	5	54	16	32	4
50%	67	12	61	30	46	10
60%	72	24	66	44	59	23
70%	77	39	70	58	63	37
80%	79	59	76	70	71	57
90%	79	73	78	75	77	73
100%	79	79	79	79	79	79

^a A statistical weight was assigned to each item by meal component with combination entree=2, condiment=0.33, remaining meal components=1.

^b Each child had 3 omission rates (one for each of 3 recalls). Omission Rate=[sum of weighted omissions/(sum of weighted omissions+sum of weighted matches)]×100. Omission rates were undefined if a child was observed to have eaten nothing regardless of what was reported eaten. Omission rates ranged from 0% to 100%, with 0% considered perfect.

^c Each child had 3 intrusion rates (one for each of 3 recalls). Intrusion Rate=[sum of weighted intrusions/(sum of weighted intrusions+sum of weighted matches)]×100. Intrusion rates were undefined if a child was observed to have eaten something yet failed to report eating anything. Intrusion rates ranged from 0% to 100%, with 0% considered perfect.

^d Each child had 3 omission rates and 3 intrusion rates (one omission rate and one intrusion rate for each of 3 recalls).

^e To examine errors of consistency in accuracy, the maximum difference across rates was tallied for each individual child's first, second, and third recall. (For example, a child with omission rates of 40%, 30%, and 50% for the first, second, and third recall would be in the 20% decile for omission rate because the maximum difference in omission rates was within 20%. A child with omission rates of 35%, 50%, and 65%, and intrusion rates of 45%, 60%, and 40% for the 3 respective recalls would be in the 30% decile for omission and intrusion rates because the maximum difference in his/her omission and intrusion rates was within 30%.)

^f To examine accuracy, the maximum rate across each individual child's first, second, and third recall was tallied. (For example, a child with omission rates of 40%, 30%, and 50% for the first, second, and third recall would be in the 50% decile for omission rate because maximum omission rate was 50%. A child with omission rates of 35%, 50%, and 65%, and intrusion rates of 45%, 60%, and 40% for the 3 respective recalls would be in the 70% decile for omission and intrusion rates because maximum omission and intrusion rates were 65%.)

Table 3

Distribution and percent of food items for matches, omissions, and intrusions by meal component across all 275 recalls from 104 children

Meal component	Match ^a	Omission ^b	Intrusion ^c	Omission Rate ^{d,e}	Intrusion rate ^{e,f}
		←No. items→			←%→
Entree	28	35	16	56	36
Combination entree	160	131	80	45	33
Bread/grain	163	150	175	48	52
Vegetable	70	176	71	72	50
Fruit	50	121	47	71	48
Dessert	61	116	23	66	27
Condiment	103	212	55	67	35
Beverage	406	241	246	37	38
Breakfast meat	9	23	7	72	44
Miscellaneous ^g	4	33	5	89	56
Total	1,054	1,238	725	54	41

^a Match=item reported eaten and observed eaten.

^b Omission=item not reported eaten but observed eaten.

^c Intrusion=item reported eaten but not observed eaten.

^d Omission rate=[sum of omissions/(sum of omissions+sum of matches)]×100. Omission rates were undefined if a child was observed to have eaten nothing regardless of what was reported eaten. Omission rates ranged from 0% to 100%, with 0% considered perfect.

^e Statistical weights were not necessary because rates were calculated within each meal component.

^f Intrusion rate=[sum of intrusions/(sum of intrusions+sum of matches)]×100. Intrusion rates were undefined if a child was observed to have eaten something yet failed to report eating anything. Intrusion rates ranged from 0% to 100%, with 0% considered perfect.

^g Miscellaneous=foods such as candy and chips that did not fit into any of the other meal components.