Public Health Nutrition: 10(10), 1057–1066 doi: 10.1017/S1368980007683888

Validation-study conclusions from dietary reports by fourth-grade children observed eating school meals are generalisable to dietary reports by comparable children not observed

Albert F Smith[1,*], Suzanne Domel Baxter[2], James W Hardin[3], Caroline H Guinn[2], Julie A Royer[2] and Mark S Litaker[4]
[1]Department of Psychology, Cleveland State University, 2121 Euclid Avenue, Cleveland, OH 44115, USA: [2]Institute for Families in Society, University of South Carolina, Columbia, SC, USA: [3]Center for Health Services and Policy Research, and Department of Epidemiology and Biostatistics, University of South Carolina, Columbia, SC, USA: [4]Department of Diagnostic Sciences, University of Alabama at Birmingham, Birmingham, AL, USA

Submitted 14 June 2006: Accepted 7 December 2006: First published online 2 March 2007

Abstract

Objective: To investigate the effect of observing school meals on children's dietary reports.
Subjects and setting: One hundred and twenty children randomly selected, but with half girls, from usual school-meal eaters among 312 volunteers (from all 443 fourth-grade children in six schools in one district).
Design: Children were assigned randomly to one of 12 conditions yielded by crossing observation status (observed; not observed), target period (previous day; prior 24 hours), and interview time (morning; afternoon; evening).
Results: Response variables included interview length, number of meals and snacks reported for the target period, and, for two school meals, number of meal components reported, importance-weighted number of items reported and kilocalories reported. These variables were transformed to principal components; two were retained (1, the school meal variables; 2, interview length and number of meals and snacks). Analyses of variance on principal component scores tested effects of observation status, target period, interview time and all interactions. Observation status did not affect scores on either retained principal component. Scores on Component 2 showed that more intake was reported in prior-24-hours interviews than in previous-day interviews.
Conclusions: The effect of target period on reported intake indicates that the response variables were sufficiently reliable to detect manipulations. This, together with the finding that response variables did not depend on observation status, suggests that observation of school meals does not affect fourth-grade children's dietary reports, and that conclusions about dietary reports by fourth-grade children observed eating school meals in validation studies may be generalised to dietary reports by comparable children not observed.

Keywords
Validation
24-Hour dietary reports
Children
Methodology
Observation

Validation studies of dietary self-report methods are conducted to evaluate how accurately individuals report their dietary intake. In such studies, self-reported information is compared with reference information that is presumed to represent – more or less accurately – the truth about dietary intake. For example, 24-hour dietary reports (24hDRs) are compared with observations; responses to food-frequency questionnaires are compared with respondent-kept food records; nutrient-transformed food-frequency responses are compared with nutrient-transformed 24hDRs; and energy- and nutrient-transformed 24hDRs are compared with biomarkers (e.g. references 1–6).

Evaluating reporting accuracy is the primary concern of any validation study. However, as results of validation studies are used to estimate the reporting accuracy of individuals for whom there is no reference information, another central concern should be the generalisability of conclusions about response validity to such individuals[7–10].

Although some methods of collecting reference information about individuals' dietary intake are unlikely to influence their reports, others may. Specifically, record-keeping by individuals about their own food intake, or being observed while they eat, may influence their subsequent reports about food intake. Any effect of a

*Corresponding author: Email a.f.smith@csuohio.edu

validation procedure on self-reports would compromise the generalisability of conclusions about the accuracy of self-reports.

In studying the dietary-reporting accuracy of fourth-grade children (who are approximately 10 years old), we have obtained reference information by observing children eating school meals[2,11–17]. Moreover, observation of school meals has been used to obtain reference information in most validation studies of children's unassisted dietary reports (see reference 18 for a review of publications between 1970 and 1999; for examples, see references 2, 7, 11–17, 19–25).

Such observation may affect the behaviour of the observed children: it may influence what they eat; it may enhance their attention to what they eat so that they subsequently report more accurately; it may lead them to report more carefully. If so, what is learned about the responses of observed children, including the accuracy of those responses, would not be generalisable to the responses of children who had not been observed. However, in practice, it is the responses of children who have not been observed in which investigators are interested, because survey respondents and research participants have typically not been observed (e.g. references 26–29).

The ideal validation procedure would have no effect on the reports of individuals about whom reference information is collected[9,10]; ideally, to establish that a validation procedure does not affect children's dietary reports, one would establish the statistical equivalence[30–33] of the reports of children who are and who are not subjected to that procedure on some set of variables measured on the reports. As we will elaborate in the Discussion, establishing statistical equivalence is potentially very costly, with cost increasing with the stringency of the equivalence criterion. This leads to consideration of alternative strategies, one of which we undertook in the present work.

To investigate the effect of observing fourth-grade children on their subsequent dietary reports, we compared 24hDRs of children who were observed eating school meals with reports of children who were not observed. We analysed five variables that might be influenced by observation; these variables did not include reporting accuracy, because observation is required to assess reporting accuracy. Finding no statistically significant effect of observation status would encourage the conclusion that observation does not affect children's dietary reporting performance, but the absence of a statistically significant effect does not imply statistical equivalence[31]: a true difference might not be detected either because of insufficient power or because the measured variables are psychometrically unreliable. To address the latter concern, we evaluated whether the five variables were affected by a manipulation of the particular 24-hour target period about which children were to report: in their 24hDRs, children were to report about either the prior 24 hours (from 24 hours before the interview until the time the interview began) or the previous day (from midnight to midnight). These conditions differed in the amount of time between the target period and the time of the interview, which plausibly affects how much intake children report. Finding that, regardless of observation status, the response variables are affected by a manipulation (e.g. target period) would indicate that the measured variables were sensitive and reliable, and would add support to the conclusion that school meal observation does not affect fourth-grade children's dietary reports[9,34].

Methods

The Institutional Review Board of the Medical College of Georgia approved all aspects of the study.

Participants

The participants were 120 fourth-grade children recruited from six of 33 elementary schools in one school district. These schools were selected to obtain a sample of children with high participation in school meals; at these schools, 70% (from 58% to 82%) of children were eligible to receive free or reduced-price school meals during the school year of data collection.

During autumn 2002, all fourth-grade children from these schools were invited to participate. The invitation explained that participating children might be watched while eating breakfast and lunch at school several times, and might be interviewed about what they eat; that not all children who agreed to participate would be interviewed; and that a $US 15 cheque would be mailed for each interview. Of the 443 children (29% black boys, 30% black girls, 20% white boys, 16% white girls, 3% other boys, 2% other girls) invited to participate, 312 (70%) agreed by returning signed child assent and parental consent forms. (Agreement rates did not depend on race or sex[2].) From these 312, the 120 participants (39 black boys, 45 black girls, 18 white boys, 13 white girls, three other boys, two other girls) were randomly selected, subject to the constraints that half be girls and that all be usual eaters of school breakfast and school lunch. A child satisfied the latter criterion by eating both school meals on at least four of five practice observation days.

Design and procedure

Within sex, children were assigned randomly to be observed or not observed, and then to one of the six interview conditions yielded by crossing two target periods (previous day; prior 24 hours) with three interview times (morning; afternoon; evening). Thus, 10 children (five girls) participated in each of the 12 observation status $\times$ target period $\times$ interview time conditions. The distribution of children by race was essentially identical in the two observation-status conditions (observed: 42 black, 16 white, two other; not observed: 42 black, 15 white, three other).

For the previous-day target period, children were to report about midnight to midnight of the day before the interview, beginning with when they got up (e.g. for an interview on Tuesday at 6.30 pm, from midnight Sunday to midnight Monday). For the prior-24-hours target period, children were to report about the period from 24 hours before the interview until the time the interview began (e.g. for an interview on Tuesday at 6.30 pm, from 6.30 pm on Monday to 6.30 pm on Tuesday) (see reference 2). Available evidence indicates that fourth-grade children's dietary reporting accuracy does not depend significantly on whether interviews are conducted in person or by telephone[15], so morning and afternoon interviews were conducted in person at school after breakfast and lunch, respectively, and evening interviews were conducted by telephone after 6.30 pm.

Interviewers were blind to the observation status of the children they interviewed: observations and interviews were conducted by different individuals; interviewers did not know at which schools observations were conducted for specific meals.

We have published analyses of the effects of target period and interview time on the reporting accuracy of the 60 observed children[2].

Observations

For each observation, one of two dietitians stood, for an entire meal period, by the table(s) at which classes or groups of children regularly sat, and appeared to observe the entire class or group although she recorded information about only one or two children; during such observation periods, all fourth-grade children at that school who had agreed to participate in the study wore nametags provided by the research staff. Children knew when an observer was present, but did not know who, specifically, was being observed or who, if anyone, would be interviewed. To facilitate identifying what children ate, only children who obtained meals from the school food service were observed[35]. To familiarise children with the presence of an observer and to identify usual school-meal eaters, practice observations were conducted with each class prior to data collection. Inter-observer reliability, assessed weekly using established procedures[2,13–15,36], was acceptable; observers agreed on the identities of all items and on amounts within one-fourth serving for 90% of the items.

As children were randomly assigned to be observed or not observed, observations were conducted in all six schools. However, no observations were conducted in the schools of children who were interviewed without having been observed during meal periods that occurred during the 24-hour periods about which these children were interviewed.

Interviews

Each interview was conducted by one of two dietitians using a multiple-pass protocol modelled on that of the Nutrition Data System for Research (NDS-R, version 4.05_33; Nutrition Coordinating Center, University of Minnesota). The NDS-R target period is the previous day. For interviews about the prior 24 hours, the NDS-R protocol was modified so that a child was to report intake for the interview day, from when he or she awoke until the time of the interview, followed by intake for the preceding day, from 24 hours before the interview until the time he or she went to bed. If a child reported intake prior to the beginning of the target period, the interviewer asked about the next time he or she ate or drank until report was for the target period. A fuller description of the interview protocols is given elsewhere[2].

Interviewers both recorded children's reports on paper forms and made audio-recordings. Interview quality, evaluated daily during practice sessions before data collection and during data collection on randomly selected interviews, was acceptable[2,13–15,37].

Response variables

We evaluated five variables – interview length, number of meals and snacks reported for the target period, and, for the two school meals, number of meal components reported, importance-weighted number of items reported and kilocalories reported.

For a reported meal to be counted as a school meal, a child had to indicate that the meal was eaten at school, report the time to within an hour of the scheduled time, and name the meal appropriately[2,12–15]. Every reported school-meal item was classified as one of 10 meal components – lunch entrée, combination entrée, bread/cereal, vegetable, fruit, dessert, condiment, beverage, breakfast meat, and miscellaneous. The importance weight assigned to combination entrées (e.g. hamburger) was 2; to condiments was 1/3; and to other meal components was 1. For example, a child who reported eating non-zero amounts of Cheerios, white milk and orange juice at school breakfast and a hamburger, French fries, mayonnaise, corn, chocolate milk and vanilla pudding at school lunch reported six unique meal components (bread/cereal, beverage, combination entrée, condiment, vegetable, dessert) and 9.33 importance-weighted items. We have used these methods of classifying and importance-weighting in other studies[2,12–15].

For each child, for each reported school-meal item, reported amount and kilocalories/serving information were used to compute reported kilocalories; these were summed across items reported for the two school meals. Children reported amounts in servings as none (0), taste (0.1), little bit (0.25), half (0.5), most (0.75), all (1.0), or the number of servings for more than one serving. We have quantified servings this way in other studies[2,12–15].

Kilocalorie information for standard school servings of school-meal items was obtained from the NDS-R database or from the school district's nutrition programme.

Analyses

Anticipating appreciable correlations between response variables, we planned to conduct principal component analysis of these variables; to retain a subset of the principal components, rotated orthogonally; and to conduct analysis of variance of scores on retained principal components. Principal component analysis transforms intercorrelated response variables into mutually uncorrelated principal components, each of which is a linear combination of the original variables. The first principal component accounts for maximum variability in the response variables; each successive principal component captures a maximum amount of remaining variability. Although five principal components are required to account for all variability in five response variables, a smaller number might capture a substantial fraction of the variability and be retained. To enhance interpretability, retained principal components are rotated orthogonally to maximise the correlations of some original variables with each component while minimising others. For each participant, a score on each retained principal component may be calculated; this is a linear combination of that participant's values on the original variables[38,39].

Statistical analyses were conducted with SAS (version 9.1; SAS Institute).

Results

Twelve children – six observed and six not observed – reported no meals that satisfied the criteria to be considered school meals. In practice, one would not know whether an interviewed child had eaten particular meals. Therefore, we included all children in the analyses, using, for children who reported no meals that satisfied the criteria to be considered school meals, zeroes as the values of meal components, items and kilocalories reported for the two school meals. (Analyses that excluded these children agreed essentially with those that included them; we will discuss one discrepancy not relevant to observation status.)

Table 1 shows, for each response variable, the mean and the standard deviation of the children in each of the conditions defined by observation status, target period and interview time. The correlations columns of Table 2 show the overall matrix of pairwise correlations between response variables: number of meals and snacks reported for the target period was highly correlated with interview length; and the three school meal variables (number of meal components reported, importance-weighted number of items reported and kilocalories reported) were highly intercorrelated; but pairwise correlations between variables in the former set and variables in the latter were low.

The high correlations between some variables confirmed the appropriateness of using principal component analysis to form a smaller number of composite variables. Two principal components were retained that, together, accounted for over 88% of the variability in the five response variables. The communality column of Table 2 shows the proportion of variability in each response variable accounted for by the retained principal components; each communality is at least 0.86. The loadings columns of Table 2, which show the correlations of every response variable with each retained principal component, indicate that Component 1 is, essentially, a composite of variables containing information reported for the two school meals (number of meal components, importance-weighted number of items and kilocalories) and Component 2 is, essentially, a composite of interview length and the number of meals and snacks reported for the target period.

Table 3 shows, for each of the 12 conditions, the mean and the standard deviation of scores on each retained principal component. Separate analyses of variance of scores on each component tested effects of observation status, target period, interview time and all interactions.

Whether children had been observed did not affect any analysed aspect of their reports. Table 3 shows that, averaged over target-period and interview-time conditions, mean scores on each retained principal component were similar in the two observation-status conditions (Component 1 effect size $= 0.0932$ SD, Component 2 effect size $= 0.0194$ SD; SD = standard deviation). For scores on neither retained principal component was the effect of observation status, or any interaction of observation status with other variables, significant: for Component 1 scores, all $F < 0.4$ (all $P > 0.6$); for Component 2 scores, all $F < 0.5$ (all $P > 0.6$).

We examined the effects of the target-period and interview-time manipulations. Target period should affect reporting performance: children in the previous-day condition were interviewed, on average, 25 hours after the midpoint of their target period; children in the prior-24-hours condition were interviewed 12 hours after the midpoint of their target period. Thus, children interviewed about the prior 24 hours could be expected to remember and/or report more, and therefore to report for longer, than those interviewed about the previous day.

Table 3 shows, consistent with this speculation, that Component 2 scores were, on average, higher in the prior-24-hours conditions than in the previous-day conditions, $F(1,108) = 9.68$, $P < 0.005$. Table 1 corroborates that, compared with children who reported for the previous day, children who reported for the prior 24 hours reported more meals and snacks for the target period and had longer interviews. No other effect on Component 2 scores was statistically significant.

The only significant effect on Component 1 scores, which were essentially composites of information

https://doi.org/10.1017/S1368980007683888 Published online by Cambridge University Press

Table 1 Means (and standard deviations) of variables measured in dietary reports*

	Interview length (min)		No. of meals and snacks reported for target period†		No. of meal components reported for two school meals‡,§		Importance-weighted no. of items reported for two school meals‡,¶		Kilocalories reported for two school meals‡,‖	
Target period/Interview time	Observed**	Not observed**	Observed**	Not observed**	Observed**	Not observed**	Observed**	Not observed**	Observed**	Not observed**
Prior 24 hours††										
Morning‡‡	16.7 (6.4)	16.9 (6.5)	4.2 (1.8)	4.7 (1.7)	3.2 (1.5)	3.6 (1.8)	5.5 (2.7)	5.7 (3.2)	475 (320)	585 (363)
Afternoon‡‡	17.1 (6.1)	16.9 (5.4)	4.3 (2.0)	4.7 (2.0)	4.2 (1.4)	4.4 (1.6)	7.7 (2.1)	6.3 (2.1)	760 (202)	565 (281)
Evening‡‡	16.9 (5.0)	15.2 (3.8)	4.4 (1.3)	4.1 (1.3)	3.9 (2.5)	3.1 (1.9)	5.7 (3.6)	4.1 (2.9)	490 (415)	449 (294)
Mean§§	16.6 (5.4)		4.4 (1.6)		3.7 (1.8)		5.8 (2.9)		554 (323)	
Previous day††										
Morning‡‡	14.7 (2.8)	12.3 (3.4)	4.0 (0.9)	3.6 (1.3)	3.5 (2.5)	3.3 (1.9)	5.9 (3.7)	5.5 (3.6)	570 (372)	536 (331)
Afternoon‡‡	12.3 (4.7)	12.7 (3.4)	3.6 (1.1)	3.8 (1.1)	3.2 (1.3)	3.0 (2.4)	5.3 (2.5)	4.3 (3.4)	446 (236)	457 (323)
Evening‡‡	15.0 (3.6)	14.7 (4.3)	3.9 (1.0)	4.1 (1.4)	4.6 (2.1)	4.3 (1.3)	7.1 (3.3)	7.6 (3.3)	668 (491)	872 (458)
Mean§§	13.6 (3.8)		3.8 (1.1)		3.7 (2.0)		5.9 (3.4)		592 (391)	
MEAN¶¶	15.5 (5.0)	14.8 (4.8)	4.1 (1.4)	4.2 (1.5)	3.8 (1.9)	3.6 (1.8)	6.2 (3.0)	5.6 (3.2)	568 (357)	578 (361)

* $n = 120$ fourth-grade children, with 10 children (five girls) in each of 12 conditions defined by observation status, target period and interview time. Each variable was measured in the dietary report of each child.
† Number of meals and snacks reported eaten by the child at school, at home and elsewhere during the 24 hours about which he or she was to report.
‡ School meals are breakfast and lunch prepared by school food services and eaten at school during the target period about which the child was to report.
§ Meal components include lunch entrée, combination entrée, bread/cereal, vegetable, fruit, dessert, condiment, beverage, breakfast meat, and miscellaneous.
¶ The importance weight of combination entrées was 2; of condiments was 1/3; and of items of every other meal component was 1.
‖ For each child, kilocalories reported for the two school meals were calculated by summing, over reported items, the products of amount in servings and kilocalories per serving obtained from NDS-R (Nutrition Data System for Research) or from the school district's nutrition programme.
** For observed children, school meals (breakfast and lunch) that were eaten during the target period about which the child was to report were observed by an observer who recorded what the child ate and drank; for children not observed, no such observation took place during the target period about which the child was to report, and so no reference information was recorded.
†† The prior-24-hours target period was from 24 hours before the interview until the time the interview began; the previous-day target period was from midnight to midnight of the day that preceded the interview.
‡‡ Morning and afternoon interviews were conducted in-person at school after breakfast and lunch, respectively; evening interviews were conducted by telephone after 6.30 pm.
§§ Means for each variable averaged across observation status and interview time.
¶¶ Means for each variable averaged across target period and interview time.

Table 2 Overall correlations between pairs of response variables and principal component analysis*

	Correlations						Loadings on retained orthogonally rotated principal components††	
	Interview length	No. of meals and snacks reported for target period†	No. of meal components reported for two school meals‡,§	Importance-weighted no. of items reported for two school meals‡,¶	Kilocalories reported for two school meals‡,‖	Communality**	Component 1	Component 2
Interview length	1.00					0.87	0.27	0.89
No. of meals and snacks reported for target period	0.74	1.00				0.89	0.07	0.94
No. of meal components reported for two school meals	0.37	0.21	1.00			0.86	0.92	0.15
Importance-weighted no. of items reported for two school meals	0.42	0.26	0.86	1.00		0.93	0.94	0.20
Kilocalories reported for two school meals	0.35	0.20	0.77	0.86	1.00	0.87	0.92	0.13

* $n = 120$ fourth-grade children, with 10 children (five girls) in each of 12 conditions defined by observation status, target period and interview time.
† Number of meals and snacks reported eaten by the child at school, at home and elsewhere during the 24 hours about which he or she was to report.
‡ School meals are breakfast and lunch prepared by school food services and eaten at school during the target period about which the child was to report.
§ Meal components include lunch entrée, combination entrée, bread/cereal, vegetable, fruit, dessert, condiment, beverage, breakfast meat, and miscellaneous.
¶ The importance weight of combination entrées was 2; of condiments was 1/3; and of items of every other meal component was 1.
‖ For each child, kilocalories reported for the two school meals were calculated by summing, over reported items, the products of amount in servings and kilocalories per serving obtained from NDS-R (Nutrition Data System for Research) or from the school district's nutrition programme.
** The communality is the proportion of variability of the response variable that is captured by the retained principal components.
†† Loadings are correlations between principal components and response variables.

Table 3 Means (and standard deviations) of scores on two retained principal components*

Target period†	Interview time‡	Observation status§ Observed	Not observed	Mean¶
(a) Component 1‖				
Prior 24 hours	Morning	−0.27 (0.8) [1]‡‡	−0.10 (1.0) [1]	
	Afternoon	0.47 (0.7) [0]	0.10 (0.8) [0]	−0.06 (0.9)
	Evening	−0.11 (1.2) [2]	−0.45 (1.0) [2]	
Previous day	Morning	−0.03 (1.2) [2]	−0.08 (1.0) [1]	
	Afternoon	−0.22 (0.7) [1]	−0.38 (1.1) [2]	0.06 (1.1)
	Evening	0.44 (1.3) [0]	0.63 (0.9) [0]	
Mean‖		0.05 (1.0)	−0.05 (1.0)	
(b) Component 2††				
Prior 24 hours	Morning	0.25 (1.3)	0.44 (1.1)	
	Afternoon	0.20 (1.4)	0.40 (1.4)	0.28 (1.1)
	Evening	0.32 (0.9)	0.08 (0.9)	
Previous day	Morning	−0.09 (0.5)	−0.48 (0.8)	
	Afternoon	−0.46 (0.9)	−0.32 (0.7)	−0.28 (0.7)
	Evening	−0.17 (0.7)	−0.17 (0.8)	
Mean**		0.01 (1.0)	−0.01 (1.0)	

* $n = 120$ fourth-grade children, with 10 children (five girls) in each of 12 conditions defined by observation status, target period and interview time. All tabulated means are rounded to the closest hundredth; all tabulated standard deviations are rounded to the closest tenth.
† The prior-24-hours target period was from 24 hours before the interview until the time the interview began; the previous-day target period was from midnight to midnight of the day that preceded the interview.
‡ Morning and afternoon interviews were conducted in-person at school after breakfast and lunch, respectively; evening interviews were conducted by telephone after 6.30 pm.
§ For observed children, school meals that were eaten during the target period about which the child was to report were observed by an observer who recorded what the child ate and drank; for children not observed, no such observation took place during the target period about which the child was to report, and so no reference information was recorded.
¶ Means are for each level of target period, averaged over levels of interview time and observation status.
‖ Component 1 reflects intake reported for school meals. School meals are breakfast and lunch. (Number of meal components reported for two school meals, importance-weighted number of items reported for two school meals and kilocalories reported for two school meals had high loadings on Component 1.)
** Means are for each level of observation status, averaged across target period and interview time. Effect size, measured in SD (standard deviation), is the difference between the condition means divided by the pooled standard deviation. For Component 1, the effect size for observation status is 0.0932; for Component 2, it is 0.0194.
†† Component 2 reflects interview length and the number of meals and snacks reported for the target period (the 24-hour period about which the child was to report). (These two variables had high loadings on Component 2.)
‡‡ Values in square brackets are the number of children in each condition who reported no meals that satisfied the criteria to be considered school meals.

reported about the two school meals, was the interaction of target period and interview time, $F(2,108) = 4.99$, $P < 0.01$. This effect appears to have been due to the particular distribution over the six target-period by interview-time conditions of the 12 children who reported no meals that satisfied the criteria to be considered school meals (see panel a of Table 3). This interaction was not significant in the analysis that excluded these children, and was the only discrepancy between the two analyses. In a larger study, minor variation over conditions in the distribution of such children would have less impact, and this effect would likely not be replicated.

Discussion

A fundamental question about dietary self-reports collected in epidemiological surveys and research studies is 'To what extent do the reports reflect true intake?' To answer this question, validation studies compare self-reported dietary information with reference information obtained by a method considered more valid.

In numerous validation studies of children's dietary reports, the 'more valid' method is observation of school meals[2,7,11–17,19–25]. Observing children while they eat may influence their dietary reports; more generally, obtaining reference information to assess response validity may influence self-reports[7–10,40]. The generalisability of results of validation studies hinges on whether collecting such reference information affects reports.

We assessed the effect of school meal observation on fourth-grade children's dietary reports by analysing five response variables, and found two key results. First, a composite of one subset of these variables was affected in a predictable way by a manipulated variable, target period. This indicates that the response variables were sufficiently reliable to detect manipulations. Second, we found no dependence of dietary reporting performance on observation status (or on any interaction of observation status with other manipulated variables). This suggests that observation of school meals – at least as we have implemented it – does not affect children's dietary reports. Together, these results suggest that conclusions about fourth-grade children's dietary reports from validation studies in which children are observed eating school meals may be generalised to dietary reports from comparable children who are not observed.

Limitations

One limitation of this study is that these results only suggest, but do not guarantee, that observation of school meals does not affect children's dietary reporting performance. With 60 children per observation-status condition and a two-tailed type I error rate of 0.10, the power of this experiment was 0.85 to detect a moderate effect (0.5 SD) of observation status in a conventional test of the null hypothesis of no effect. However, to have power of 0.80 to detect a small effect (0.2 SD) with a two-tailed type I error rate of 0.10 would require 310 children per observation-status condition[41,42].

On Component 1, the composite variable on which we observed the larger difference between observation-status conditions, the condition means differed by 0.0932 SD, which, as half of a conventionally defined small effect[42], is likely negligible. However, to assert that the means of the two observation-status conditions were statistically equivalent, using as the equivalence criterion an interval of 0.19 SD (which is smaller than a small effect), would require finding this difference in an experiment with 831 children per observation-status condition[31]. This sample-size requirement is sobering.

The combination of no significant effect of observation status and the significant effect of target period encourages, but does not guarantee, the conclusion that school meal observation does not affect fourth-grade children's dietary reporting performance. An open issue is how large an effect of observation status would have to be to undermine claims about the generalisability of validation-study conclusions to individuals not observed. This deserves serious attention from the community of investigators interested in validation-study methodology[30,31].

A second limitation is that we have at best established that there is not a specific effect of observation of particular school meals on children's dietary reports. Our data do not exclude the possibility of some general effect. We compared the responses of children observed at school meals during the period about which they would report with the responses of children not observed at the school meals (and during which no observers were present) during the period about which they would report. However, during the study, the latter children ate other school meals during which observers were present and so during which the children followed the study protocol (e.g. wore nametags). To exclude a general effect of school meal observation would require collecting data from children who had no exposure to observers. This would likely be impossible, given that elementary-school children eating school meals are generally monitored by adults[15,43].

A third limitation is that the participants were fourth-grade children who tended to be from economically disadvantaged families in one school district. It would be valuable to broaden the investigation of the effect of school meal observation on dietary reporting performance to children of other ages and other socio-economic levels, and to assess relationships with children's cognitive ability and scholastic achievement.

Other relevant research

There appears to be little published research on effects of school meal observation on children's dietary reporting performance. However, in evaluating a software system for eliciting dietary reports from fourth-grade children, Baranowski *et al.*[7] recognised that collecting reference information by observing school lunch might influence children's reports. From each participating child, they obtained two 24hDRs – one using the self-operated software and one in a dietitian-administered interview – and evaluated the correspondence of these 24hDRs. Correspondence measures were based on matches (items reported in the software 24hDR and to the dietitian), intrusions (items reported in the software 24hDR but not to the dietitian) and omissions (items reported to the dietitian but not in the software 24hDR). To assess the effect of observation, Baranowski *et al.* compared correspondence between 24hDRs of children observed at school lunch on the previous day with the correspondence between 24hDRs of children who had not been observed. These investigators found that observed and unobserved children did not differ on the match and omission measures, but that observed children averaged significantly lower than unobserved children on the intrusion measure. The conclusion of Baranowski *et al.*, that 'observation of school lunch somehow intensifies the experience of lunch, which impacts accuracy of report by minimizing report of foods not eaten…' (p. 384), differs from ours. However, in the study of Baranowski *et al.*, observation status was partially confounded with a 'bogus-pipeline' manipulation intended to enhance reporting accuracy. Data in their article permitted assessment of the unconfounded effect of observation status; analysis of these showed that the difference in the intrusion measure for observed and unobserved children, although in the direction discussed by Baranowski *et al.*, was not statistically significant.

Final remark

A significant research challenge for nutritional epidemiology is identifying methods that optimise the validity of dietary reports by respondents who do not anticipate reporting and who report without memory aids. This requires a research phase in which validity is assessed, and a purpose of this phase should be to estimate the response validity of individuals for whom reference information will not be available. In other words, the response validity of validation-study participants should be an estimate of the response validity of respondents in general. We have described an approach to assessing whether collecting reference information by observing school meals in a validation study with fourth-grade children influences their dietary reports. Our results suggest that it does not.

Absent a complete theory of the impact of measurement procedures on behaviour, assessing the impact of measurement in validation studies is essential to generalising the results of validation studies to research participants for whom reference information has not been collected. In addition, a consensus on the magnitude of an effect of collecting reference information that can be treated as trivial or inconsequential must be developed. These should be continuing concerns of researchers interested in the usefulness of methods used to collect dietary information.

Acknowledgements

Sources of funding: This research was supported by grant R01 HL63189 from the National Heart, Lung, and Blood Institute of the National Institutes of Health. S.D.B. was Principal Investigator.

Conflict of interest declaration: No author has any conflict of interest.

Authorship responsibilities: Study concept – A.F.S., S.D.B.; study design – S.D.B., A.F.S.; acquisition of data, recruitment of schools and participants, data coding and entry, quality control – S.D.B., C.H.G.; statistical analysis and interpretation of data – A.F.S, J.W.H., S.D.B., J.A.R., M.S.L.; drafting and revising the manuscript – all authors; obtained funding – S.D.B.

Acknowledgements: These results were presented in somewhat different form at the Second Annual Meeting of the International Society for Behavioral Nutrition and Physical Activity (Quebec City, 2003). We appreciate the cooperation of children, faculty and staff at the schools at which we collected data, and of the school district's nutrition programme and board of education. We thank Michelle L Baglio, Francesca HA Frye and Nicole M Shaffer for their assistance with data collection. We thank J Whitney Keener for her assistance in preparing the manuscript. We thank Tom Baranowski, Erin Fekete and Amy F Subar for their comments on earlier versions of this article and for their permission to acknowledge them.

References

1 Conway JM, Ingwersen LA, Vinyard BT, Moshfegh AJ. Effectiveness of the US Department of Agriculture 5-step multiple-pass method in assessing food intake in obese and nonobese women. *American Journal of Clinical Nutrition* 2003; **77**: 1171–8.
2 Baxter SD, Smith AF, Litaker MS, Guinn CH, Shaffer NM, Baglio ML, *et al*. Recency affects reporting accuracy of children's dietary recalls. *Annals of Epidemiology* 2004; **14**: 385–90.
3 Thompson FE, Subar AF, Brown CC, Smith AF, Sharbaugh CO, Jobe JB, *et al*. Cognitive research enhances the accuracy of food frequency questionnaire reports: results of an experimental validation study. *Journal of the American Dietetic Association* 2002; **102**: 212–8, 223–5.
4 Sevak L, Mangtani P, McCormack V, Bhakta D, Kassam-Khamis T, dos Santos Silva I. Validation of a food frequency questionnaire to assess macro- and micro-nutrient intake among South Asians in the United Kingdom. *European Journal of Nutrition* 2004; **43**: 160–8.
5 Heitmann BL, Lissner L. Can adverse effects of dietary fat intake be overestimated as a consequence of dietary fat underreporting? *Public Health Nutrition* 2005; **8**: 1322–7.
6 Subar AF, Kipnis V, Troiano RP, Midthune D, Schoeller DA, Bingham S, *et al*. Using biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: The OPEN Study. *American Journal of Epidemiology* 2003; **158**: 1–13.
7 Baranowski T, Islam N, Baranowski J, Cullen KW, Myres D, Marsh T, *et al*. The Food Intake Recording Software System is valid among fourth-grade children. *Journal of the American Dietetic Association* 2002; **102**: 380–5.
8 Smith AF. Cognitive processes in long-term dietary recall. *Vital and Health Statistics* 1991; **6**(4). Also available at http://www.cdc.gov/nchs/data/series/sr_06/sr06_004.pdf. Accessed 31 January 2007.
9 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.
10 Webb EJ, Campbell DT, Schwartz RD, Sechrest L. *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Chicago, IL: Rand McNally, 1966.
11 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. *Journal of the American Dietetic Association* 1997; **97**: 1293–8.
12 Baxter SD, Thompson WO, Litaker MS, Frye FHA, Guinn CH. Low accuracy and low consistency of fourth-graders' school breakfast and school lunch recalls. *Journal of the American Dietetic Association* 2002; **102**: 386–95.
13 Baxter SD, Smith AF, Guinn CH, Thompson WO, Litaker MS, Baglio ML, *et al*. Interview format influences the accuracy of children's dietary recalls validated with observations. *Nutrition Research* 2003; **23**: 1537–46.
14 Baxter SD, Thompson WO, Smith AF, Litaker MS, Yin Z, Frye FHA, *et al*. Reverse versus forward order reporting and the accuracy of fourth-graders' recalls of school breakfast and school lunch. *Preventive Medicine* 2003; **36**: 610–4.
15 Baxter SD, Thompson WO, Litaker MS, Guinn CH, Frye FHA, Baglio ML, *et al*. Accuracy of fourth-graders' dietary recalls of school breakfast and school lunch validated with observations: in-person versus telephone interviews. *Journal of Nutrition Education and Behavior* 2003; **35**: 124–34.
16 Domel SB, Thompson WO, Baranowski T, Smith AF. How children remember what they have eaten. *Journal of the American Dietetic Association* 1994; **94**: 1267–72.
17 Baxter SD, Thompson WO, Davis HC. Prompting methods affect the accuracy of children's school lunch recalls. *Journal of the American Dietetic Association* 2000; **100**: 911–8.
18 McPherson RS, Hoelscher DM, Alexander M, Scanlon KS, Serdula MK. Dietary assessment methods among school-aged children: validity and reliability. *Preventive Medicine* 2000; **31**: S11–33.
19 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. *Journal of the American Dietetic Association* 1994; **94**: 626–30.
20 Lytle LA, Murray DM, Perry CL, Eldridge AL. Validating fourth-grade students' self-report of dietary intake: results from the 5-A-Day Power Plus program. *Journal of the American Dietetic Association* 1998; **98**: 570–2.

21 Lytle LA, Nichaman MZ, Obarzanek E, Glovsky E, Montgomery D, Nicklas T, *et al*. Validation of 24-hour recalls assisted by food records in third-grade children. *Journal of the American Dietetic Association* 1993; **93**: 1431–6.

22 Samuelson G. An epidemiological study of child health and nutrition in a northern Swedish county. II. Methodological study of the recall technique. *Nutrition & Metabolism* 1970; **12**: 321–40.

23 Todd KS, Kretsch MJ. Accuracy of the self-reported dietary recall of new immigrant and refugee children. *Nutrition Research* 1986; **6**: 1031–43.

24 Warren JM, Henry CJK, Livingstone MBE, Lightowler HJ, Bradshaw SM, Perwaiz S. How well do children aged 5–7 years recall food eaten at school lunch? *Public Health Nutrition* 2003; **6**: 41–7.

25 Weber JL, Lytle L, Gittelsohn J, Cunningham-Sabo L, Heller K, Anliker JA, *et al*. Validity of self-reported dietary intake at school meals by American Indian children: The Pathways Study. *Journal of the American Dietetic Association* 2004; **104**: 746–52.

26 Moreira P, Padez C, Mourao I, Rosado V. Dietary calcium and body mass index in Portuguese children. *European Journal of Clinical Nutrition* 2005; **59**: 861–7.

27 Wiley AS. Does milk make children grow? Relationships between milk consumption and height in NHANES 1999–2002. *American Journal of Human Biology* 2005; **17**: 425–41.

28 Nicklas TA, Morales M, Linares A, Yang S-J, Baranowski T, De Moor C, *et al*. Children's meal patterns have changed over a 21-year period: The Bogalusa Heart Study. *Journal of the American Dietetic Association* 2004; **104**: 753–61.

29 Champagne CM, Bogle ML, McGee BB, Yadrick K, Allen HR, Kramer TR, *et al*. Dietary intake in the lower Mississippi delta region: results from the Foods of our Delta Study. *Journal of the American Dietetic Association* 2004; **104**: 199–207.

30 Rogers JL, Howard KI, Vessey JT. Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin* 1993; **113**: 553–65.

31 Tryon WT. Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: an integrated alternative method of conducting null hypothesis statistical tests. *Psychological Methods* 2001; **6**: 371–86.

32 Seaman MA, Serlin RC. Equivalence confidence intervals for two-group comparisons of means. *Psychological Methods* 1998; **3**: 403–11.

33 Feinstein AR. *Principles of Medical Statistics*. Boca Raton, FL: Chapman & Hall/CRC Press, 2002.

34 Frick RW. Accepting the null hypothesis. *Memory & Cognition* 1995; **23**: 132–8.

35 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. *Journal of the American Dietetic Association* 1992; **92**: 219–21.

36 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. *Journal of the American Dietetic Association* 2004; **104**: 1385–93.

37 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. *Journal of the American Dietetic Association* 2004; **104**: 1577–85.

38 Hatcher L, Stepanski EJ. *A Step-by-step Approach to Using the SAS System for Univariate and Multivariate Statistics*. Cary, NC: SAS Institute, 1994; Ch. 14.

39 SAS Institute. *SAS/STAT User's Guide, Version 8*. Cary, NC: SAS Institute, 1999; Ch. 6, 26 and 52.

40 Rebro SM, Patterson RE, Kristal AR, Cheney CL. The effect of keeping food records on eating patterns. *Journal of the American Dietetic Association* 1998; **98**: 1163–5.

41 Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ: Erlbaum, 1988.

42 Cohen J. A power primer. *Psychological Bulletin* 1992; **112**: 155–9.

43 Simons-Morton BG, Baranowski T. Observation in assessment of children's dietary practices. *Journal of School Health* 1991; **61**: 198–200.