



E-FAN-01-013

December 2001

# Designs for Measuring How the School Breakfast Program Affects Learning

Prepared by Jonathan Jacobson, Ronette Briefel, Philip Gleason,  
and Rachel Sullivan of Mathematica Policy Research, Inc.

ERS contact: Jayachandran N. Variyam

## Abstract

This report describes a study design permitting a scientifically defensible evaluation of the impact of the School Breakfast Program (SBP) on learning and cognitive development among children. Following presentation of a literature review and conceptual framework of the SBP-learning relationship, four alternative designs for measuring this relationship are proposed and assessed, including: (1) the design for an experimental evaluation of the impact of the universal-free SBP; (2) an experimental design involving random assignment of classrooms into the SBP among schools applying to participate in the program for the first time; (3) a nonexperimental design involving analysis of Early Childhood Longitudinal Study (ECLS) data; and (4) a nonexperimental design involving analysis of forthcoming National Health and Nutrition Examination Survey (NHANES) data. Among these alternatives, the ECLS-based design (with supplemental analysis of 1988-1994 NHANES III data) was selected as most promising, and the report presents further details concerning this design.

## Acknowledgments

The authors thank Jay Variyam, ERS project officer; Mary Begalle, Rob Hollister, and Rob Meyer, members of the expert panel; Michael Murphy and Nicholas Zill, consultants to MPR; John Endahl and Anita Singh, Food and Nutrition Service, USDA; and ERS staff members David Smallwood, Elizabeth Frazao, Margaret Andrews, Mark Prell, and Josh Winicki, as well as other participants in the May 2000 expert panel meeting.

**This report was prepared by Mathematica Policy Research, Inc., under a research contract from the Economic Research Service. The views expressed are those of Mathematica and not necessarily those of ERS or USDA.**

# Contents

Chapter	Page
Executive Summary .....	vi
I. Introduction .....	1
II. Literature Review .....	3
A. The Literature .....	3
1. The link between nutrition and cognitive development.....	3
2. The link between breakfast and cognition .....	6
3. The effects of the School Breakfast Program.....	8
B. Implications of Research.....	17
1. Conceptual framework.....	17
2. Design Alternatives.....	22
III. Alternative Designs.....	25
A. Key Elements of the Alternative Designs .....	25
B. The USBP Design .....	27
1. Intervention and counterfactual.....	27
2. Basic design approach.....	28
3. Data collection .....	31
4. Measurement of key characteristics and outcomes .....	34
5. Analysis plans .....	36
6. Statistical power.....	39
7. Design costs .....	40
8. Design summary: strengths and weaknesses.....	41
C. SBP Applicant Design.....	43
1. Intervention and counterfactual.....	43
2. Basic design approach.....	44
3. Data collection .....	46
4. Measurement of key characteristics and outcomes .....	47
5. Analysis plans .....	47
6. Statistical power.....	49

7.	Design costs .....	51
8.	Design summary: strengths and weaknesses.....	52
D.	ECLS-K-based Design .....	53
1.	Intervention and counterfactual.....	53
2.	Basic design approach.....	54
3.	Data collection .....	56
4.	Measurement of key characteristics and outcomes .....	56
5.	Analysis plans .....	57
6.	Statistical power.....	60
7.	Design costs .....	60
8.	Design summary: strengths and weaknesses.....	61
E.	NHANES Design .....	61
1.	Intervention and counterfactual.....	62
2.	Basic design approach.....	62
3.	Data collection .....	64
4.	Measurement of key characteristics and outcomes .....	69
5.	Analysis plans .....	70
6.	Statistical power.....	71
7.	Design costs .....	72
8.	Design summary: strengths and weaknesses.....	73
F.	Assessing the Alternative Designs .....	74
1.	Summary of alternative designs .....	74
2.	Choosing the most feasible design.....	79
IV.	ECLS-K-NHANES III Design .....	82
A.	ECLS-K Nonexperimental Comparison Group Design .....	82
B.	Variables in the ECLS-K.....	83
1.	Outcome variables.....	83
2.	Participation in the SBP .....	85
3.	Background characteristics of students and schools .....	88

C.	ECLS-K Supplemental Data Collection.....	92
1.	Parent survey.....	92
2.	School administrator survey.....	94
D.	ECLS-K Analytic Approach .....	97
1.	Selection bias .....	98
2.	Propensity score methods.....	100
3.	Minimum detectable differences.....	101
E.	Presentation of ECLS-K-Based Design Results.....	103
F.	Supplemental NHANES III Analysis.....	105
1.	Basic design approach.....	105
2.	Data and measurement .....	106
3.	Previous research .....	113
4.	Analysis plans .....	114
5.	Summary.....	116
	References .....	117
	Appendix A .....	124
	Appendix B .....	130
	Appendix C .....	136
	Appendix D.....	138

## Tables

<b>Table</b>	<b>Page</b>
II.1 Summary of studies of the effect of SBP/USBP participation on school outcomes .....	11
II.2 Implications of conceptual model with respect to intermediate outcomes.....	20
III.1 SBP applicant design minimum detectable differences .....	50
III.2 Expected sample sizes for NHANES in years 2000, 2001, and 2002.....	63
III.3 Current variables in NHANES pertinent to school breakfast and learning.....	65
III.4 Minimum detectable differences on achievement test scores .....	72
III.5 Summary of evaluation design options .....	75
III.6 Strengths and weaknesses of evaluation design options .....	76
IV.1 Student background characteristics in the ECLS-K.....	88
IV.2 School background characteristics in the ECLS-K (in addition to average levels of student characteristics) .....	91
IV.3 Supplemental survey questions for parents.....	93
IV.4 Supplemental survey questions for school administrators.....	95
IV.5 Presentation of estimated impacts of the SBP on a given outcome (for example, test scores measured in standard deviation units).....	104
IV.6 Sample size of children aged 6 to 16 years, by SBP participation.....	109
IV.7 Sample sizes of children aged 6 to 16 years, by SBP participation and income level.....	109
IV.8 Sample sizes of children aged 6 to 16 years, by SBP participation and dietary and cognitive test score data .....	110
IV.9 Sample sizes of children aged 6 to 16 years, SBP participation and dietary and cognitive test score data .....	111

## Figure

Figure	Page
II.1 Conceptual framework for analyzing the effects of SBP participation on learning .....	19

## **Executive Summary**

This report documents the development of alternative designs to study how participation in the School Breakfast Program (SBP) influences learning. The project, conducted by Mathematica Policy Research, Inc. (MPR) for the Economic Research Service of the U.S. Department of Agriculture (USDA), began with a review of the relevant literature, followed by the development of several alternative designs for studying the relationship between SBP participation and learning. The most feasible alternative was expanded into a detailed study design and analysis plan. The layout of this report mirrors the progression of the project work, from the literature review and description of the alternative designs to a detailed presentation of the most feasible one.

### **Justification for the Project**

A thorough understanding of the nature of the relationship between SBP participation and cognitive performance would provide useful information for school administrators and policymakers. Administrators of schools that do not participate in the SBP could use the information to help decide whether to offer school breakfasts, and policymakers could use it to decide whether and how to promote school participation in the program. Although a separate evaluation of the universal-free school breakfast program (USBP) is underway, information about the effects of the regular SBP on learning might help policymakers determine the appropriate direction for SBP policy.

Previous research shows that eating breakfast improves performance on short-run cognitive tests (see Briefel et al., 1999, in Chapter II, for literature reviews). The research also shows that SBP participation both increases the intake of selected vitamins and minerals and improves school attendance (Devaney and Fraker 1989; and Murphy et al., 1998a and 1998b). Although these findings suggest that eating a school breakfast improves learning, studies that have estimated this relationship directly have failed to produce definitive results. No study has been able to conclude convincingly that eating a school breakfast leads to improvements in long-term or short-term cognition or academic achievement. The inconclusiveness of these studies reflects their limitations, such as the use of small samples, unreliable measures of SBP participation and of academic achievement, or the failure to control adequately for selection effects in the context of nonexperimental designs.

The designs proposed in this report build on the evidence in the literature. They assume an underlying conceptual model in which participation in the SBP directly influences dietary intake and school attendance—intermediate outcomes that ultimately may lead to improved learning. Eating which has the potential to improve their nutritional and health status. Improved nutritional and health status, in turn, are hypothesized to influence three key outcomes: (1) short-term cognition, (2) brain development and functioning, and (3) behavior in school. Health status may influence school attendance as well. All these factors may contribute to learning, the ultimate outcome of interest, which typically is measured by students' grades or performance on standardized academic or achievement tests.

## Alternative Designs

MPR proposed four alternative designs that might be used to study the effect of SBP participation on learning: (1) an experimental evaluation of the USBP pilot project (the USBP design), (2) an experimental study of the effects of classroom implementation of the SBP on breakfast program applicant schools (the SBP applicant design), (3) a nonexperimental study of the effects of the SBP based on data from the Early Childhood Longitudinal Study, Kindergarten Cohort (the ECLS-K design), and (4) a nonexperimental study of the effects of the SBP based on data from the forthcoming National Health and Nutrition Examination Survey (the NHANES design).

**USBP Design.** The first alternative is based on design work that MPR has conducted under contract to USDA's Food and Nutrition Service. This design would use an experimental approach to examine the effects on student outcomes of USBP pilot programs funded by USDA and operating in elementary schools in six school districts. The outcomes would include dietary intake and academic achievement. Under this design, the elementary schools in each district would be matched to one another on the basis of their characteristics and then randomly assigned to either a treatment group that would offer the USBP or to a control group that would continue using the regular SBP. After implementation of the USBP in the treatment schools, samples of students would be drawn from each school. Survey data and administrative data would provide information on students, and outcomes of students in USBP schools would be compared with outcomes of students in control (regular SBP) schools to generate an estimate of the impact of USBP *availability*.

A major strength of this design is its experimental approach, which would generate unbiased estimates of the impact of USBP availability on student outcomes. In addition, USDA has funded this evaluation, which Abt Associates Inc., currently is conducting. The design's weaknesses include its expected high costs (\$13.0–\$13.3 million, including demonstration costs) and limited statistical power. It would not directly yield experimentally based (and unbiased) estimates of the impact of participating in the USBP or of eating a free school breakfast.

**SBP Applicant Design.** No previous study of a U.S. breakfast program other than the ongoing USBP study has used an experimental design to estimate the effects of participation on learning/academic achievement. Although a study of this type would be difficult to implement, it is important to fully explore the benefits and feasibility of using an experimental design. One potentially feasible experimental design is the SBP applicant design, which would involve random assignment of elementary school classrooms in schools applying to participate in the SBP for the first time. Rather than serve to permanently approve or deny SBP participation to the applicant schools' treatment and control classrooms, random assignment would be a mechanism to either enable classrooms to begin serving school breakfasts immediately (treatment status) or delay the SBP until a subsequent year (control status). During the first year, treatment classrooms would serve school breakfasts and control classrooms would not. The breakfasts would be served free to all students to encourage schools to participate in the demonstration, and to encourage students to eat school breakfasts. This design would yield experimental and unbiased estimates of the impact of the offer of free school breakfasts (as opposed to no school breakfasts) on student outcomes measured by survey and administrative data.

The SBP applicant design would have the benefits of random assignment and internal validity. However, because it probably would have to be implemented in a limited geographic area (or set of areas), it also would likely have only limited external validity. Also, there are certain aspects associated with the design, such as randomizing classrooms and serving breakfast in the classroom, which could cause problems when actually implemented. In addition, the design would be difficult to implement, and its cost would be high (\$10.7–\$11.3 million, which includes the costs of the demonstration, data collection, analysis, and report writing). Finally, given the similarity between the SBP applicant design and the USBP design, it is unclear what new information could be derived from this design.

***ECLS-K-Based Design.*** The ECLS-K is a longitudinal dataset with information on a large, nationally representative sample of students from the kindergarten class of 1998. The dataset will ultimately include follow-up information collected through the students' fifth grade year (if they progress at a normal rate in school) and includes a wealth of information on their cognitive development, academic progress, and SBP participation status. This design alternative proposes using ECLS-K data in a nonexperimental examination of the effects of SBP participation on learning. Under this design, measured outcomes in SBP participants and in nonparticipants would be compared, after controlling to the extent possible for observed (and potentially unobserved) characteristics of the two groups.

The strengths of the ECLS-K design are that it would be nationally representative and would be based on large sample sizes of students and schools, giving the design substantial statistical power. The wealth of information in the ECLS-K dataset would allow the design to be based on a range of high-quality measures of student achievement and health status (height, weight, etc.) at various points in time. A final strength of the design is its relatively low estimated cost (\$0.4–\$0.5 million). A major limitation is that it would be nonexperimental, so that estimates of the impact of SBP participation on learning could be subject to selection bias. Although the design includes plans for addressing the selection bias issue, these methods are imperfect. Another weakness of the design is that it would not include estimates of dietary intake as an intermediate outcome.

***NHANES-Based Design.*** The NHANES provides data on a nationally representative sample of students. The current NHANES data collection effort began in 1999 with a series of continuous, annual surveys. The NHANES-based design would use the current NHANES data to generate nonexperimental estimates of the impact of SBP participation on student outcomes. This design is similar in its basic approach to the ECLS-K design. However, it differs in that it focuses on dietary intake, nutritional status, and health status, all of which are intermediate outcomes that may ultimately be related to learning. Because the current NHANES dataset does not include the various measures of student achievement available in the ECLS-K, it would not be possible to use the NHANES-based design to measure the effect of SBP participation on student achievement unless a major supplemental data collection effort were undertaken. The data collection effort would be costly (contributing \$2.9–\$3.0 million in addition to the \$1.0–\$1.1- million analysis and reporting costs) and would involve administration of achievement and cognitive tests to students.

The major strength of the NHANES-based design is its ability to examine outcomes relating to dietary intake, nutritional status, and health status. Much like the ECLS-K design, this design would

produce nationally representative results. Its main weakness is its nonexperimental approach. Finally, because the sample size of the annual NHANES survey is much smaller than that of the ECLS-K, several years of NHANES data would have to be used to generate sufficient statistical power.

### **Assessing the Alternative Designs**

A key factor in the choice of the most promising design is the fact that USDA has implemented the USBP design. The status of the USBP design does not preclude us from recommending that design as most promising one, and, by extension, from suggesting that no other design be considered. However, it does imply that any other design must be recommended in the context of the ongoing study of the USBP. We therefore rule out recommending the SBP applicant design, as it is similar to the USBP design and would yield little new information. Moreover, its implementation would be very difficult and costly. The two other designs are nonexperimental. We believe that the additional cost and number of years it would take to gather an adequate sample prevent us from assessing the NHANES-based design as being the most promising of the two.

We consider a modified version of the ECLS-K design as the most promising alternative. The ECLS-K design addresses at least two limitations of the USBP design, as it (1) would be nationally representative, and (2) would have large samples of students and schools, which would yield substantial statistical power. Furthermore, the rich information on students' achievement and cognitive functioning would enable us to address the key outcomes of interest in the study. The nonexperimental nature of the ECLS-K design is mitigated by the fact that the USBP design also examines the relationship between the breakfast program and learning, thereby providing a second methodological perspective on the issue. Another limitation of the ECLS-K design is its inability to obtain information on intermediate outcomes, such as nutritional status and health status. However, it is possible to correct this weakness by adding a supplemental analysis component to the design. In particular, analysis of the NHANES III dataset, which is similar to the current NHANES data collection effort but was collected during the period from 1988 through 1994 (and includes measures of students' cognition/achievement), would enable the design to examine the impact of SBP participation on both intermediate outcomes and the ultimate outcomes of interest. We refer to this combined set of analyses as the ECLS-K-NHANES III design.

### **The ECLS-K-NHANES III Design**

SBP participation, defined as usually eating a school breakfast provided through a USDA-funded SBP, is the intervention to be examined by the ECLS-K-NHANES III design. The counterfactual aspect as defined in the design, is *not* usually eating a school breakfast. The design calls for estimating the differences in outcomes between students who participate in the SBP and those who do not participate. The study design, which is nonexperimental, would incorporate two analyses. The primary analysis would be based on ECLS-K data; the supplemental analysis would be based on data from NHANES III. The design would be primarily cross-sectional, with learning-related outcomes and SBP participation status measured at a given point in time (although with some longitudinal analysis of ECLS-K data). The design's nonexperimental nature would create the risk of selection bias, but the use of instrumental variables (IV) methods would address this problem.

***ECLS-K Data and Measurement Issues.*** The ECLS-K is important because it is a nationally representative, longitudinal survey that is large enough to support the analysis of the relationship between SBP participation and learning in the overall sample and among key subgroups. The database contains information on three types of outcomes: (1) cognitive functioning, (2) events and processes associated with learning, and (3) other aspects of a child’s growth. This variety of outcomes permits exploration of many of the possible effects of SBP participation. Although the information on school breakfast consumption contained in the dataset is not sufficiently precise for accurate measurement of the number of school breakfasts a child eats per week across the full school year, it could be used to create a binary or categorical variable that would measure usual SBP participation. Finally, the ECLS-K includes a rich set of variables that measure the background characteristics of students and schools. These variables are important because they could be used to control for factors that influence SBP participation as well as learning.

In addition to the existing ECLS-K data, the design calls for the use of supplemental data from the parent and administrator components of the ECLS-K survey. Suggested supplemental questions on the parent survey relate to children’s breakfast habits; parents’ attitudes about breakfast; and children’s morning schedules, including their commute to school. Suggested supplemental questions on the school administrator survey would obtain information on the number of eligible and participating children in the SBP and National School Lunch Program (NSLP), by type of payment level (free, reduced-price, full-price); the schools’ morning schedules; and the schools’ reasons for not participating in the SBP. The supplemental data collection would provide useful descriptive information and could be used in the estimation of instrumental variable (IV) models to account for selection bias.

***Analytic Approach.*** Because the proposed ECLS-K analysis would be nonexperimental, multivariate statistical methods would have to be used to infer the impact of school breakfast participation (or, alternatively, attendance at a school offering the SBP) on the educational outcomes of otherwise identical students. In the most basic set of models, ordinary least squares (OLS), fixed (school)-effects, or random-effects estimation is suggested. However, the strategy for estimating the impact of SBP participation on learning also should account for the possible selection of SBP participants into the sample, conditional on students attending a school that offers the SBP. In addition, it may be necessary to account for the selection of schools into the sample of schools offering the SBP. Selection of either type, if based on unobserved factors correlated with learning outcomes, would bias OLS estimates of the impact of SBP participation on learning.

An IV procedure is a standard method of correcting for selection bias. IV models require the selection of instrumental, or identifying, variables that are uncorrelated with the outcome variable, but that are correlated with the endogenous explanatory variable—in this case, SBP participation status (and, perhaps, SBP availability at a school). Collecting supplemental data from parents and school administrators would increase the number of candidates for instrumental variables in the analysis. The key instrumental variable proposed for this design would be a variable constructed from information collected in both the regular ECLS-K data collection and through the supplemental data collection designed to measure the timing of children’s arrival at school, the start of the school day, and the children’s mode of transportation to school.

An alternative estimation strategy is based on propensity score modeling. Under this approach, probabilities of SBP participation by students in SBP schools would be estimated, making it possible to create subgroups of likely SBP participants in SBP schools and likely SBP participants in non-SBP schools. Such a *propensity score-based* strategy could help in estimating the impact of *offering* the SBP to students, as well as the impact of SBP *participation* on learning.

Each proposed estimation method has limitations. A simple comparison of reported differences in outcomes between SBP participants and SBP nonparticipants would not account for observed differences between the two groups of students. OLS, fixed (school)-effects, and random-effects estimates would account for observed differences between students and their schools, but not for the endogenous selection of students into the SBP (and, possibly, for selection into schools offering the SBP). IV estimates might help correct for sample selection bias, but the precision of the resulting impact estimates is likely to depend on the predictive power of the instrumental variables used in the analysis. Propensity score methods could account for any association between attendance at an SBP school and different outcomes for students with different probabilities of SBP participation. However, the ability of propensity score methods to identify the actual impact of SBP participation on learning outcomes depends on whether attendance at an SBP school is independent of unobserved factors influencing learning.

***Supplemental NHANES III Analysis.*** Adding NHANES III data analysis related to SBP participation and learning to the ECLS-K data would enhance what could be inferred from the ECLS-K alone about the relationship between eating a school breakfast and learning. The NHANES III is a nationally representative database that contains information collected from 1988 to 1994 on school-aged children’s school breakfast participation, dietary intakes, nutritional status, and health status (intermediate outcomes in our design), and on their performance on standardized cognitive tests. The survey also includes information on explanatory, or control, variables such as family characteristics; on school outcomes, such as attendance and grade level; and on behavioral outcomes.

The NHANES III data provide a unique opportunity to fully examine the relationships between SBP participation and such intermediate outcomes as nutrition and health status; this examination could not be performed with the ECLS-K dataset. In addition, NHANES III offers a cross-sectional examination of the relationship between school breakfast participation and two key indicators of learning—cognitive functioning and academic performance—in a national sample of participating and nonparticipating children. Although the ECLS-K and other national surveys also measure much of the same family and school data, NHANES III includes detailed information on diet, nutrition, and health pertaining to school-aged children.

The supplemental NHANES III analysis would provide additional evidence on the impact of SBP participation on learning and may corroborate (or may fail to corroborate) the primary findings from the ECLS-K. However, although the NHANES III includes a wealth of information pertinent to the research question of interest, the total sample size is smaller than that of the ECLS-K, and the dataset is not longitudinal. The cross-sectional nature of the NHANES III dataset (and the resulting NHANES III analysis) means that key outcomes are measured at a given point in time in the school year. This feature of the NHANES III dataset would limit the extent of an analysis of academic achievement, as it does not measure year-to-year change. It also would reduce the statistical power

of NHANES-based impact estimates of SBP participation on learning. Nevertheless, the additional evidence from the NHANES III on the research question still would be valuable because the issues of selection bias and statistical power are important to the ECLS-K-based design.

To estimate the effects of SBP participation on learning, regression analysis would be used to compare cognitive outcomes of participants and nonparticipants, after controlling for relevant, measurable factors. In comparing academic and learning outcomes in participants and nonparticipants, the design calls for models that account for such important factors as food insufficiency and poor nutritional status, as well as factors that relate to prenatal nutrition, such as low birth weight and exposure to cigarette smoke. Environmental exposures, such as to lead, revealed by elevated blood lead levels, should also be considered in interpreting the results of cognitive tests and academic performance. Thus, the set of controls available in the NHANES III data will be richer in some ways than those available in ECLS-K. Unlike the ECLS-K, however, the NHANES III dataset includes few school-level control variables (or indicators of which schools sample members attend). Thus, this analysis would not measure or control for some of the institutional factors associated with the schools that students attend.

## **Conclusion**

A complete understanding of the relationship between eating a school breakfast and learning is very important to policymakers and school administrators who must make decisions about the direction of the SBP. The designs described in this report, particularly the ECLS-K-NHANES III design, could be used to perform analyses that would significantly add to the existing research on the impacts of SBP participation on outcomes related to learning. This design also has the advantage of maximizing the utility of existing, ongoing, national surveys of U.S. schoolchildren through supplemental ECLS-K data collection and analysis as well as analysis of NHANES III data. Finally, the ECLS-K-NHANES III design would complement the current USBP evaluation, which uses an experimental approach to study a similar set of issues.