


IMPROVING GOVERNMENT PERFORMANCE

**Joseph S. Wholey,
Kathryn E. Newcomer,
and Associates**

*Evaluation
Strategies
for
Strengthening
Public
Agencies
and Programs*

Next Page: M.J. WARGO Article

Characteristics of Successful Program Evaluations



Michael J. Wargo

This chapter describes and analyzes three unusually successful evaluations of national food and nutrition programs.* These exemplary evaluations focused on the Food Stamp Program, the National School Lunch and Breakfast programs, and the Supplemental Food Program for Women, Infants, and Children (more commonly referred to as WIC). All three evaluations were judged to be successful on the basis of their demonstrated impact on one or more of the primary functions of the agency responsible for their administration—the Food and Nutrition Service (FNS) of the U.S. Department of Agriculture.

First, the context within which these evaluations operated, their objectives, methods, and results, and their impact on the FNS are discussed. Then, characteristics common to these three dissimilar program evaluations that contributed to their significant impact on program and agency operations are identified. Finally, it is suggested that if the characteristics that contributed to the success of these evaluations are replicated they will contribute to the utilization of any evaluation.

The three evaluations discussed in this chapter focus on the FNS's primary programs, which together have an annual appropriation that exceeds \$17 billion. By far the largest of these programs

*The views expressed in this chapter are those of the author and should not be construed to be those of either the U.S. Department of Agriculture or the U.S. General Accounting Office.

is the Food Stamp Program, with an annual budget of \$12 billion. The annual appropriation for the National School Lunch and Breakfast Program exceeds \$3.7 billion and that of the WIC is approximately \$1.6 billion. The combined budgets of these three programs represent approximately 40 percent of the annual budget of the U.S. Department of Agriculture.

The Food Stamp Program provides food vouchers to families with incomes below the official U.S. poverty income guidelines. The National School Lunch and Breakfast Program provides funds and agricultural commodities to school systems that provide nutritionally adequate meals to school-age children. The price charged students for school breakfast and lunch is based on measures of their family's economic need. WIC provides food supplementation in the form of food packages or food vouchers to low-income pregnant, breastfeeding, and postpartum women as well as to their infants and children to age five. Although all of these programs are designed to provide food supplementation, they differ significantly in terms of size, target population, eligibility requirements, form of food supplementation, and mechanism for and organizations involved in service delivery (U.S. Department of Agriculture, 1983).

As indicated earlier, the three evaluations that are the subject of this analysis were judged unusually successful in terms of their impact on a national food program or on the agency responsible for their administration (FNS). More specifically, these three evaluations were judged successful because they had a direct impact on one or more of the following products of agency function: legislative initiatives, regulations, budget, policy, operations, and program performance. Although the specific impacts associated with these three successful evaluations differed, they all, in some way, had a positive impact on agency functions and program operations.

Three Successful Evaluations

A recently completed national evaluation of eight work registration models for the Food Stamp Program fits the previously mentioned criteria for success (Brandeis University and Abt Associates, 1986). This study assessed the effectiveness of eight

models de
food stam
participar
tions, and
eight mo
twelve-mo
througho
individual
to the pr
procedure
and partic
period. Th
sense that
treatment
longitudin
participan

The
strated tha
a variety o
were found
stamp ben
benefits th

The
had a signi
legislation
ment and
contribute
follow to
the model
technical ;
Food Stamp
impact on
assistance,

And
focused on
Wish a
This eval
school dist

models designed to facilitate finding employment for able-bodied food stamp recipients. Program impacts were measures in terms of participant employment and earnings, food stamp benefit reductions, and the relationship of benefits accrued to program costs. The eight models were fully implemented and then evaluated over a twelve-month period at eighteen sites geographically dispersed throughout the United States. At each site, program-eligible individuals were randomly assigned to a work registration model or to the program's standard work registration and job search procedures (control condition). Data were collected from managers and participants at the eighteen sites throughout the twelve-month period. This study can be characterized as a true experiment in the sense that there was random assignment of participants to the treatment (model) and control (standard) conditions. It also was longitudinal in nature in that data were repeatedly collected on participants over eleven months.

The Food Stamp Work Registration Evaluation demonstrated that all eight models could be successfully implemented at a variety of sites. All of the work registration and job search models were found to increase employment and earnings and to reduce food stamp benefit payments. All but one of the eight models produced benefits that exceeded their costs.

The results of the Food Stamp Work Registration Evaluation had a significant impact on the development of new food stamp legislation that required all states to develop food stamp employment and training programs by April 1, 1987. The evaluation also contributed to the development of regulations that states must follow to implement the new legislation. Finally, descriptions of the models used in the demonstration have been included in technical assistance material disseminated to states. In short, the Food Stamp Work Registration Evaluation has had a significant impact on food stamp legislation, regulations, policy, technical assistance, program operations, and service delivery.

Another unusual/ successful evaluation conducted by FNS focused on the National School Lunch and Breakfast programs (Wellish and others, 1983a, 1983b; Radzikowski and Gale, 1984). This evaluation was designed to describe the characteristics of school districts, schools, and students participating in the National

School Lunch and Breakfast programs; to determine program impacts on the nutritional intake of participants and on their family's food expenditures; and to identify the determinants of program participation.

The National Evaluation of School Nutrition Programs collected data from a nationally representative sample of public school districts, schools, students, and their families. Program participation, dietary intake, and growth and development data were collected from approximately 6,500 elementary and secondary school students. Data on family composition, income and food expenditures were obtained from the students' parents. School district, school, and program characteristic data were collected by mail survey from principals and food service managers in 1,100 public schools. These data were synthesized and then analyzed in terms of the study's primary objectives.

The results of this evaluation indicated that the National School Lunch Program was available to 98 percent of the students in public elementary and secondary schools. In contrast, the National School Breakfast Program was available to only 39 percent of those students. Participation in both programs was highest for students from families with incomes in the lowest quartile of total income. Both the lunch and the breakfast programs supplemented family food expenditures; that is, families did not reduce food expenditures when their children received subsidized meals at school. The school lunch was found to be nutritionally superior to noon meals eaten by nonparticipants, and the total twenty-four-hour nutritional intake of participants was determined to be superior to that of nonparticipants. The study found that the primary determinant of participation in the lunch program was the price charged for the meals.

The National Evaluation of School Nutrition Programs has had a significant impact on the FNS's policy, legislative, and budget formulation processes. The study produced a very large and extensive participant characteristic data base as well as a model for estimating program participation. Both of these products have proven invaluable in the preparation of budget proposals for the programs and in the development of policy and legislative alternatives. The results and conclusions of the evaluation have also

Characteristics of

been used to justify
als and to develop
quality of meals
Program.

The third e
study that focus
of women, infants
criteria are used to
and nutritional ris
of program on pr
and their infants a
185 percent of the
be certified by a he
program's regulati
ing these criteria. C
of the actual num
tion in the program

The primary
to develop a meth
istics and number
determining the ne
resource allocation
needy individuals.

The WIC E
model that permit
(which provided e
mation) and nutriti
representative samples
U.S. Department of
1980 National Hea
1980 National Nat
Prior to developi
eligibility were bas
known to be overes
information on nu
information to acc
WIC Eligibility Eva
statewide estimate

ams; to determine program of participants and on their identify the determinants of

School Nutrition Programs representative sample of public and their families. Program growth and development data 00 elementary and secondary position, income, and food e students' parents. School ristic data were collected by d service managers in 1,100 esized and then analyzed in

indicated that the National to 98 percent of the students v schools. In contrast, the s available to only 39 percent h programs was highest for i the lowest quartile of total fast programs supplemented milies did not reduce food ceived subsidized meals at be nutritionally superior to and the total twenty-four- nts was determined to be The study found that the i the lunch program was the

ool Nutrition Programs has s's policy, legislative, and y produced a very large and base as well as a model for th of these products have of budget proposals for the of policy and legislative , of the evaluation have also

been used to justify administration budgets and legislative proposals and to develop new regulations to improve the nutritional quality of meals provided by the National School Breakfast Program.

The third example of an unusually successful evaluation is a study that focused on determining the number and characteristics of women, infants, and children eligible for the WIC. Three basic criteria are used to determine WIC eligibility: categorical, income, and nutritional risk criteria. Categorical criteria relate to the focus of program on pregnant, breastfeeding, and postpartum women, and their infants and children. Applicants' incomes may not exceed 185 percent of the federal poverty guidelines and applicants must be certified by a health professional to be at "nutritional risk." The program's regulations allow states considerable latitude in applying these criteria. Consequently, FNS has never had good estimates of the actual number of individuals who are eligible for participation in the program.

The primary objective of the WIC Eligibility Evaluation was to develop a methodology and model for estimating the characteristics and number of WIC eligibles. Such information is useful in determining the need for WIC services in each state, in developing resource allocation strategies, and in targeting services to the most needy individuals.

The WIC Eligibility Evaluation developed a mathematical model that permitted a linkage between 1980 U.S. census data (which provided excellent income and family characteristic information) and nutritional risk data derived from large nationally representative samples of women, infants, and children included in two U.S. Department of Health and Human Services surveys: the 1976-1980 National Health and Nutrition Examination Survey and the 1980 National Natality Survey National Fetal Mortality Survey. Prior to development of this model, FNS estimates of WIC eligibility were based entirely on census data. Those estimates were known to be overestimates because census data do not include any information on nutritional risk nor do they provide sufficient information to accurately determine categorical eligibility. The WIC Eligibility Evaluation provided the agency's first national and statewide estimates of WIC eligibility that included improved

categorical criteria as well as a standardized measure of nutritional risk for all participant groups.

The primary products of the WIC Eligibility Evaluation were a comprehensive set of baseline estimates of the WIC-eligible population, a model that can be used to estimate program eligibility under a variety of criteria, and a projection model for producing regular annual updates of the WIC-eligible population.

The findings of the WIC Eligibility Evaluation are currently being used to develop policies to improve and coordinate programs targeting on the most needy, to assist in developing guidance to standardize nutritional risk assessment across states, and to improve WIC resource allocation. The eligibility estimating model has also been used to explore various policy and legislative options for the WIC.

These three evaluations have been described in some detail to illustrate how successful evaluations can vary considerably in terms of their focus, methods, and results. The programs assessed varied significantly in terms of size, target populations (families, school children, pregnant women and their preschool children), eligibility requirements, form of food supplementations (vouchers, meals, and food packages); and service delivery structure, mechanism, and organization.

Evaluation methods varied from a national demonstration employing a true experimental design and longitudinal data collection through cross-sectional data collection from nationally representative samples, to a reanalysis and synthesis of existing data. Finally, the results of these successful evaluations were viewed by the various stakeholders as positive, negative, or neutral. What all three evaluations had in common was their significant impact on the operations and performance of the Food and Nutrition Service and its programs.

Characteristics Common to Successful Program Evaluations

The preceding text emphasized the differences among the three successful evaluations to demonstrate that the utility of an evaluation is not necessarily tied to its focus, methods, or results. Rather, there are attributes common to successful evaluations that

transcend success. This point is the basis for the associated methodology. Planning this change is a significant finding.

Planning

associated utilization

Table

Planning

Early identification and prioritization of audiences

Assessment and identification of audience needs

Early stakeholder participation and participation

Stakeholder involvement in evaluation

nutritional

Evaluation
AIC-eligible
in eligibility
r producing
t.

are currently
ite programs
guidance to
d to improve
odel has also
otions for the

1 some detail
nsiderably in
rams assessed
ons (families,
ool children),
ms (vouchers,
cture, mecha-

emonstration
itudinal data
om nationally
sis of existing
ns were viewed
neutral. What
ificant impact
and Nutrition

Evaluations

ces among the
ie utility of an
ods, or results.
valuations that

transcend those characteristics and directly contribute to their success. Analysis of the three successful evaluations substantiates this point by indicating these evaluations had common characteristics that were related to their success and that appear to be more associated with the evaluation process than with evaluation focus, method, or results. Further, these common evaluation characteristics can be neatly categorized into three evaluations phases, namely, planning, conduct, and dissemination of results. The remainder of this chapter describes characteristics of exemplary evaluations that significantly contributed to the ultimate utilization of their findings.

Planning

All three successful evaluations described had characteristics associated with their planning process that contributed to their utilization (see Table 2). For the purposes of this discussion,

Table 2. Characteristics Common to Successful Program Evaluation.

Planning	Evaluation	
	Conduct	Dissemination
Early identification and prioritization of audiences	Formation of Evaluation Teams that include representatives of key stakeholder groups	Aggressive dissemination policy
Assessment and prioritization of audience information needs	Evaluation Team monitoring of evaluation plan implementation	Stakeholder participation in the development and implementation of the dissemination plan
Early stakeholder identification and selection for participation	Evaluation Team participation in data analysis and report writing	Timely dissemination of evaluation results to key audiences
Stakeholder involvement in evaluation design	Production of technically sound information on schedule and within budget Tailoring of reports to key audience information needs	

evaluation planning includes all activities associated with the development of objectives, an evaluation design, a data collection and synthesis plan, a data analysis plan, and data collection instruments. In short, the planning process includes all evaluation design and development efforts from the determination of the need for the evaluation up to implementation of the final evaluation design.

1. Early Audience Identification, Needs Assessment, and Needs Prioritization. In addition to meeting the success criteria, all three evaluations devoted considerable effort to identification of the evaluation's primary audiences. Further, all three evaluations in one way or another determined how important it was to meet the needs of each audience group. Then, information needs of key audience groups were assessed and information requirements for the evaluation were prioritized.

2. Stakeholder Identification and Participation Throughout Evaluation. In all three evaluations, the major stakeholders were involved early in the evaluation process. Major interested groups were identified, and their representatives were asked to participate in the evaluation design. Stakeholders assisted in translating key audience information needs into evaluation objectives and in developing a sound evaluation design that would produce the desired products on time and within budget. This active stakeholder participation was maintained throughout the planning process.

Clearly, early primary audience identification, needs assessment, and prioritization of information needs in conjunction with stakeholder participation throughout the planning process significantly contributed to the ultimate success of these evaluations.

Conduct

Characteristics associated with the conduct of these evaluations appeared to contribute significantly to their usefulness (see Table 2). For the purposes of this discussion, the conduct phase of the evaluation begins with implementation, continues through the final evaluation design, and concludes with the final report, which summarizes the study's objectives, methods, results, conclusions, and recommendations.

3. Formu

evaluations, tea
conduct the eval
implementation
participated in d

Evaluation
nenced agency pr
for the success o
team generally w
specialists, as we
affect program
external staff obt
technical, substan
team and to form
were integrated v
role. They were, l
to the direction o
evaluation team l

In all three
by at least one ad
state and local p
group represent
experts such as w
specialists, pediat
ists, and nutrition
qualified experts

These adv
major stakeholder

4. Evaluat

mentation. Anot
these three evalu
plan implementa
in the application
group, evaluation
evaluation. Probl
were developed b
ensuring that th
schedule and with

activities associated with the design, a data collection and data collection instrument. All evaluation design includes all evaluation design and analysis plans and participation of the need for the final evaluation design.

1. Needs Assessment, and ing the success criteria, all effort to identification of the all three evaluations in important it was to meet the information needs of key information requirements for

1. Participation Through- s, the major stakeholders process. Major interested representatives were asked to Stakeholders assisted in a needs into evaluation evaluation design that would and within budget. This maintained throughout the

identification, needs assessment needs in conjunction with planning process significant of these evaluations.

conduct of these evaluations to their usefulness (see ion, the conduct phase of on, continues through the th the final report, which ods, results, conclusions.

3. Formulation of Evaluation Teams. In all three exemplary evaluations, teams with unique characteristics were formed to conduct the evaluations. These evaluation teams closely monitored implementation of the evaluation design and analysis plans and participated in development of the final report.

Evaluation teams were led by highly qualified and experienced agency program evaluators who were ultimately responsible for the success of the evaluation. The members of the evaluation team generally were agency policy analysis and program operation specialists, as well as experts from other department agencies that affect program operations. In all three cases, consultants and external staff obtained under contract were used to supplement the technical, substantive, and managerial expertise of the evaluation team and to form field data collection staffs as needed. Contract staff were integrated within the evaluation team and played an equal role. They were, however, as were the other team members, subject to the direction of the agency's evaluation expert designated as the evaluation team leader.

In all three cases, these evaluation teams were also supported by at least one advisory panel. Advisory panels were composed of state and local program managers, special interest and advocacy group representatives, evaluation specialists, and substantive experts such as welfare and agricultural economists, public health specialists, pediatricians, child development experts, food specialists, and nutritionists. Advisory groups were composed of the most qualified experts who could be obtained in each desired category.

These advisory groups also included representatives of all major stakeholders.

4. Evaluation Team Monitoring of Evaluation Plan Implementation. Another characteristic associated with the success of these three evaluations was very close monitoring of evaluation plan implementation by the entire evaluation team. This resulted in the application of program management, advocacy and interest group, evaluation, and substantive expertise throughout the evaluation. Problems were anticipated and corrective action plans were developed by the team, thus avoiding catastrophic failure and ensuring that the evaluation produced the desired products on schedule and within budget.

5. Evaluation Team Participation in Analysis and Report Preparation. The entire evaluation team participated in data analysis and development of the final report. Evaluation results were interpreted by the team, who then developed the conclusions and recommendations.

6. Tailoring of Report to Various Audiences. The evaluation teams associated with these successful evaluations were responsible for the development of one or more final reports tailored to the information needs of the primary audiences—that is, the administration, Congress, and state and local program managers. Tailoring is selecting the format, content, and exposition style most appropriate for the specific audience. The purpose was to maximize the usefulness of the report to its user.

7. Dissemination. The dissemination process begins with development of a plan for distribution of findings and usually concludes when all major stakeholders are satisfied with their understanding of the results and conclusions. All three exemplary evaluations were subject to the same aggressive dissemination policy that is associated with FNS evaluation activity (Wargo, 1981). Also common to these three evaluations was stakeholder participation in development of the dissemination plan. That aggressive dissemination policy combined with stakeholder participation in development of the dissemination plan seemed to contribute significantly to the utility of these successful evaluations.

8. Aggressive Report Dissemination Policy and Timely Report Distribution. FNS's dissemination policy encourages the tailoring of evaluation reports to the needs of primary audiences, providing desired information to key decision makers in time for inclusion in the decision process and using oral presentation to supplement written reports for senior decision makers in both the executive and legislative branches of government. The primary objective of this dissemination policy is to provide needed information in a format that facilitates its use in time for its incorporation in the decision making process.

To comply with FNS's dissemination policy, all three successful evaluations developed final reports that clearly and concisely described their objectives, methodology, findings, major conclusions, and recommendations. Supplementing the final report

was a sh
primary a
objective
methodol
generaliza
itself.

Th
dum on p
to indicate
as the maj
Implicatio
guidance,
random i
departme
distributed

For
executive
agency, d
provided t
congressio
over, or in
major pro
executive s
Commerce
as the lea
sponsored

In a
and Congr
the evalua
decision-m

An
fostering o
and their
cates for d
most team
tended to p
cate when

Analysis and Report
participated in data
rt. Evaluation results
veloped the conclusions

Audiences. The evalua-
ul evaluations were
r more final reports
ry audiences—that is,
ocal program manag-
, and exposition style
The purpose was to
er.

process begins with
findings and usually
satisfied with their
. All three exemplary
essive dissemination
ctivity (Wargo, 1981).
takeholder participa-
lan. That aggressive
lder participation in
emed to contribute
valuations.

Policy and Timely
olicy encourages the
f primary audiences.
1 makers in time for
oral presentation to
1 makers in both the
ment. The primary
ide needed informa-
for its incorporation

n policy, all three
ts that clearly and
gy, findings, major
ting the final report

was a short executive summary written in the parlance of its primary audience. Each executive summary focused on the study's objectives, results, and conclusions, only briefly discussing methodology. Technical support for the validity, reliability, and generalizability of the findings was provided in the final report itself.

The basic dissemination package also included a memorandum on policy implications. The purpose of this memorandum is to indicate what the in-house portion of Evaluation Team considers as the major policy implications of the study for the administration. Implications for policy, regulations, legislative change, program guidance, and program improvement are highlighted. The memorandum is intended to promote action by senior agency and department officials. It is an internal document and is not distributed outside the department.

For all three successful evaluations, final reports and executive summaries were widely disseminated throughout the agency, department, and administration. Reports were also provided to the chairmen and ranking minority members of all congressional committees and subcommittees that have jurisdiction over, or interest in, the programs evaluated. Groups representing major program stakeholders were also sent copies of each study's executive summary. Finally, copies were sent to the Department of Commerce's National Technical Information Service, which serves as the lead agency for public dissemination of government-sponsored research.

In addition, key decision makers within the administration and Congress were briefed orally. These briefings emphasized how the evaluation's results could be utilized to improve various decision-making processes.

An important by-product of the evaluation process was the fostering of team member advocacy for the evaluation, its findings, and their dissemination. Team members became unofficial advocates for dissemination and use of the evaluation results. Further, most team members, as representatives of various stakeholders, tended to play the role of change agent as well as evaluation advocate when they returned to their respective groups (see Table 2).

Conclusion

On the basis of this analysis of three unusually successful program evaluations, it appears that the primary factors that contributed to their usefulness to key administration and congressional decision makers were (1) the establishment of evaluation teams that included the necessary technical and substantive expertise as well as representation from primary evaluation and program stakeholder groups, (2) the participation of evaluation team members throughout the evaluation process; and (3) the aggressive implementation of dissemination plans developed with stakeholder participation. Factors such as evaluation focus, content, methodology, and specific results seem to play a secondary role in determining the utilization of study findings.

References

- Brandeis University and Abt Associates, Inc. *Food Stamp Work Registration and Job Search Demonstration: Final Report*. Waltham, Mass.: Brandeis University, 1986.
- Radzikowski, J., and Gale, S. "The National Evaluation of School Nutrition Programs: Conclusions." *American Journal of Clinical Nutrition*, 1984, 40 (2), Supplement.
- U.S. Department of Agriculture. *Facts About Food Programs of the U.S. Department of Agriculture*. Washington, D.C.: U.S. Department of Agriculture, 1983.
- Wargo, M. J. "Evaluation of Nutrition Programs." In L. Datta (ed.), *Evaluation in Change: Meeting New Government Needs*. Beverly Hills, Calif.: Sage, 1981.
- Wellish, J. B., and others. *The National Evaluation of School Nutrition Programs: Final Report. Vol. 1: Overview and Presentation of Findings*. Santa Monica, Calif.: System Development Corp., 1983a.
- Wellish, J. B., and others. *The National Evaluation of School Nutrition Programs: Final Report. Vol. 2: Technical Appendices*. Santa Monica, Calif.: System Development Corp., 1983b.

Chapter

Using
Evaluation

James B.

Often a change
a program rec
sponsors at the
products of hig
rationale but sl
products will br

The feder
that must be ma
political suppo
public goods and
local level, and t
starting and char
of ineffective pro
even greater. Wh
public resources &

This chap
term evaluations
techniques are ill
month short-term
assistance—but th
redesign was inten
a short overview
section discusses h
ones that increase