META-EVALUATION: ASSESSING ALTERNATIVE METHODS OF PERFORMANCE EVALUATION AND AUDITS OF PLANNED AND IMPLEMENTED MARKETING STRATEGIES

Arch G. Woodside and Marcia Y. Sakai

ABSTRACT

A meta-evaluation is an assessment of evaluation practices. Meta-evaluations include assessments of validity and usefulness of two or more studies that focus on the same issues. Every performance audit is grounded explicitly or implicitly in one or more theories of program evaluation. A deep understanding of alternative theories of program evaluation is helpful to gain clarity about sound auditing practices. We present a review of several theories of program evaluation.

This study includes a meta-evaluation of seven government audits on the efficiency and effectiveness of tourism departments and programs. The seven tourism-marketing performance audits are program evaluations for: Missouri, North Carolina, Tennessee, Minnesota, Australia, and two for Hawaii. The majority of these audits are negative performance assessments.
Similarly, although these audits are more useful than none at all, the central conclusion of the meta-evaluation is that most of these audit reports are inadequate assessments. These audits are too limited in the issues examined; not sufficiently grounded in relevant evaluation theory and practice; and fail to include recommendations, that if implemented, would result in substantial increases in performance.

**INTRODUCTION TO META-EVALUATION**

In 1969 New York established the first state auditing and evaluation unit. A total of 61 such government departments exist in the U.S. – at least one in each of the 50 state legislatures in the United States (Brooks, 1997). Each of these auditing offices is assigned by their state legislature with the responsibility for conducting financial audits and performance audits of government departments and their specific programs. The mandate of these auditing offices is to provide answers to questions, including the following issues:

- is the audited department spending funds legally and properly in accordance to its legislative mandate, is the department’s accounting and internal control systems adequate, are the department’s financial statements accurate?
- is the department managing its operations efficiently?
- is the department achieving substantial impact in effectively accomplishing its goals?

Thus, the auditing work done for state legislative branches include two major categories of audits: (a) financial audits; and (b) performance audits. Some state audit manuals distinguish among program, operations, and management audits (e.g. *The Auditor*, 1994). For example, “a program audit focuses on how effectively a set of activities achieves objectives. A program audit can stand-alone or be combined with an operations audit. An operations audit focuses on the efficiency and economy with which an agency conducts its operations. In Hawaii the term management audit is used often to refer to an audit that combines aspects of program and operations audit. A management audit examines the effectiveness of a program or the efficiency of an agency in implementing the program or both” (*The Auditor, 1994*, pp. 1–2). In this study we use “management audit” and “performance audit” interchangeably.

The paper before you includes a meta-evaluation of available performance audits done by legislative audit offices in evaluating government tourism-marketing efforts. The term, “meta-evaluation” was created by Scriven (1969) to mean an evaluation of evaluations. We conducted a meta-evaluation on the performance audits of
tourism-marketing programs completed for the state of Tennessee (1995), Missouri (1996), North Carolina (1989), Minnesota (1985), Hawaii (1987, 1993), and for Australia (1993). These performance audits address two central issues: (1) how well are the government tourism-marketing programs being managed; and (2) how effectively are the actions of the state tourism offices contributing to their goals.

Our meta-evaluation results in two main conclusions. First, the tourism-marketing performance audits spotlight serious problems in the performances of government tourism-marketing programs. Second, the audits themselves have major shortcomings; for example, they fail to include comprehensive reporting on topics relevant in performance auditing. With one exception (i.e. Hawaii, 1987), the audit reports are not grounded in relevant theory and empirical literature on best practices. Consequently, the recommendations in the reports are too limited in scope – while beneficial, implementation of the recommendations is unlikely to have a major impact on increasing the effectiveness of state government sponsored tourism-marketing programs.

META-EVALUATION OBJECTIVES

Meta-evaluations have three central objectives. First, meta-evaluations are syntheses of the findings and inferences of research on performance – both on the managing of programs and on the effectiveness of achieving goals of programs. Thus, meta-evaluations enlighten; they help increase our knowledge and insight about what works well and poorly in managing programs.

Second, meta-evaluations are reports on the validity and usefulness of evaluation methods. Meta-evaluations include guidance in the methods useful to apply for evaluating.

Third, meta-evaluations may provide strong inference on the impact, payback, and repercussions of enacting specific decisions. Consequently, the findings in meta-evaluations help to justify and increase the confidence of legislative members and program managers, in designs and implementation of specific decisions. This third objective meets the instrumental use criterion for evaluations described by Cook (1997), “Would those who pay for evaluation be satisfied if it creates enlightenment but did not feed more directly into specific decisions? I’m not sure they would.” Related to a travel decision, Campbell (1969) provides a detailed meta-evaluation that fulfills this third objective in summarizing multiple studies on the impacts (e.g. reduction in the number of deaths) of the legislated requirement to wear safety helmets by motorcyclists.

We adopt these three objectives – enlightenment, method usefulness, and instrumental use – in preparing this report. Thus, presenting bad news is not our focus; our
focus is on increasing useful “sensemaking” (Weick, 1995) and reducing “knowing what isn’t so” (Gilovich, 1991) in evaluations. Sensemaking is creating, examining, and revising plausible explanations of events that have occurred; sensemaking is always retrospective: “People can know what they are doing only after they have done it” (Weick, 1995, p. 24). Auditing is one category of sensemaking.

Most retrospective processes used by humans cause problems: “People are extraordinarily good at ad hoc explanation” (Gilovich, 1991, p. 21). However, humans are biased strongly in favor of misinterpreting incomplete and unrepresentative data. “A fundamental difficulty with effective policy evaluation is that we rarely get to observe what would have happened if the policy had not been put into effect [implemented]. Policies are not implemented as controlled experiments but as concerted actions” (Gilovich, 1991, pp. 41–42). Without training, we rarely make and implement plans to find evidence disconfirming our beliefs. This observation applies to government tourism-marketing programs. Without collecting information about the attitudes and behavior toward the brand (e.g. Hawaii, North Carolina, or Australia) of targeted travelers not exposed to the marketing program, the success rate is evaluated – the awareness, attitudes, visits, and purchases among targeted travelers exposed to the marketing program – a research process inadequate to do the job of evaluating marketing performance.

In the second section we present a brief review of the theory-related literature on performance audits and evaluation research. Based on the conclusions drawn from the literature review, the third section offers propositions for planning and implementing a meta-evaluation of performance audits. The fourth section describes the method for collecting and analyzing performance audits of state tourism-marketing programs. Sections 5 through 11 present the findings from a meta-evaluation of the state performance audits. Section 12 provides conclusions and implications for research and practice on performance auditing of tourism-marketing programs, as well as effective strategies for managing such programs, are provided. Section 13 advocates five golden rules to apply for effective performance auditing of tourism-marketing programs.

PERFORMANCE AUDITS AND EVALUATION RESEARCH THEORY

The U.S. Office of the Comptroller General defines performance auditing as an objective and systematic examination of evidence for the purpose of providing an independent assessment of the performance of a government organization, program, activity, or function in order to provide information to improve public accountability and facilitate decision-making by parties with responsibilities
Assessing Alternative Methods of Performance Evaluation

553

to oversee or initiate corrective action. An audit is “inherently retrospective, concerned with detection of errors past – whereas many evaluative techniques can be applied retrospectively, concurrently, or prospectively” (Pollitt & Summa, 1997, p. 89).

While additional differences between audits and evaluations have been described (see Pollitt & Summa, 1997), the Association of Governmental Accountants emphasize, “Policy makers [want] reliable facts and sound, independent professional judgment, and they care little about . . . terminology. They use terms like performance auditing and program evaluation interchangeably. Their greatest concern is that they get answers to their most pressing questions about the performance of government programs and agencies” (AGA Task Force Report, 1993, p. 13, quoted from Brooks, 1997, p. 115). Thus, although recognizing that some differences exist, we use the terms “performance auditing” and “program evaluation” interchangeably.

Importance of Program Evaluation Theory

What are the key issues related to theories of performance auditing? In the case of government evaluations of state tourism-marketing programs, we believe the key issues include the following points: (1) recognition of the importance of theory in program evaluation; (2) learning the existing alternative theories of program evaluation; and (3) adopting a multiple perspective approach to program evaluation – use of multiple theories as foundations for tourism-marketing program evaluations.

Shadish, Cook and Leviton (1991, p. 20) emphasize that, “It is . . . a serious mistake to overlook the importance of theory in program evaluation.” Theory is defined as a set of assumptions used for sensemaking (i.e. a “mental model” description of some topic, see Senge, 1990). “Theory connotes a body of knowledge that organizes, categorizes, describes, predicts, explains, and otherwise aids in understanding and controlling a topic” (Shadish et al., 1991, p. 30).

Without its unique theories, program evaluation would be just a set of loosely conglomerated researchers with principal allegiances to diverse disciplines, seeking to apply social science methods to studying social programs. Program evaluation is more than this, more than applied methodology. Program evaluators are slowly developing a unique body of knowledge that differentiates evaluation from other specialties while corroborating its standing among them (Shadish et al., 1991, p. 31).

The fundamental reason theory is important for program evaluation in that the assumptions in the theory used are the rationale for the focus and method applied for the evaluation. All performance audits are based on implicit or explicit program evaluation theory. Program evaluation theory answers the questions of whether or
not the sponsored program should achieve outcomes that would not have occurred without the program, and should the program be judged by whether or not certain policies are implemented.

For the seven performance audits of government tourism-marketing programs examined in this report, we find little explication of program evaluation theory selected for use, and we find a poor match to features found in “good theory for social program evaluation” (see Chapter 2 in Shadish et al., 1991). The majority of these reports do not explicitly address the issues of program evaluation theory beyond making a general statement early in the audits. Here is an example of one such statement:

The objectives of the audit were to review the department’s legislative mandate and the extent to which the department has carried out that mandate efficiently and effectively and to make recommendations that might result in more efficient and effective accomplishment of the department’s legislative mandates . . . . The audit was conducted in accordance with generally accepted government auditing standards and includes: (1) review of applicable statutes, regulations, policies, and procedures; (2) examination of document files, contracts, data, and reports, including information compiled by the U.S. Travel Data Center; (3) interviews with a contracted vendor, directors of regional tourist associations, and staff of the Departments of Tourist Development, Transportation, Health, and Environment and Conservation; and (4) site visits to the 12 welcome centers (State of Tennessee, 1995, pp. i–ii).

These statements fail to display knowledge and understanding of alternative theories of program evaluation; and the statements fail to indicate features of good program evaluation theory. Steps toward good program evaluation theory include detailed explicit statements regarding the objectives and methods applied in a performance audit, including the auditor’s view on what constitutes a high quality evaluation of a tourism-marketing program.

For example, do the auditors agree with the views of Stake and his colleagues (1997) that “Programs designed to contribute to improvement in social well-being are to some degree meritorious just by existing.” Or, do the auditors embrace Scriven’s theory that a program’s merit should be judged by the impact of its outcomes in comparison to other programs and written standards (evaluation checklists, see Scriven, 1995). Stake et al. (1997) make clear that these views can not be held simultaneously, “…we object to Michael Scriven’s claim that the basic logic of evaluation is criterial [sic] and standards-based.”

Because of such disagreements among both program evaluation theorists and practitioners, describing a detailed, explicit, statement of the theory of program evaluation being used by a performance audit team is a valuable step. Without such a detailed explicit statement, the different stakeholders affected by the audit are likely to assume different meanings for concept words such as, “efficient” and “effective.”
Fundamental Issues in Development of Program Evaluation Theory

Related to program evaluation, fundamental issues of theory development and use include these questions:

1. Social programming: what are the important problems this program addresses? Can the program be improved? Is this program worth improving? If not, what is worth doing (including the basic issue: should a state government be involved at all in tourism-marketing, see Bonham & Mak, 1996)?

2. Knowledge use: how can I make sure my results get used quickly to help this program?

3. Valuing: is this a good program? By which notion of “good”? What justifies this conclusion?

4. Knowledge construction: how do I know the information is accurate? What counts as a confident answer? What causes that confidence?

5. Evaluation practice: given my limited skills, time, and resources, and given the seemingly unlimited possibilities, how can I narrow my options to do a feasible evaluation? What is my role – educator, methodological expert, and/or judge – of program worth? What questions should I ask, and what methods should I use? (adapted from Shadish et al., 1991, p. 35).

Several well-formed alternative theories of evaluation have different – and often opposing – answers to these five issues. The details of these theories cannot be covered in depth here; however, we briefly describe the main assumptions held by a leading proponent of each theory.

To create a learning aid, we oversimplify these theories using two dimensions: valuing results vs. valuing activity, and logical (post)positivism vs. relativism. (See Fig. 1.) The specific placement of theorists in Fig. 1 is arbitrary to some extent, but does reflect the kernel assumptions of the cited theorists. The kernel theoretical propositions of each theorist are described below.

The two dimensions found in Fig. 1 address the issues: (1) what should be valued in program evaluation; and (2) what method should be used for doing evaluations? Valuing results is performance auditing built on the proposition that a (state tourism-marketing) program should achieve certain outcomes that would not have occurred from other programs – including having no program. Valuing activities/processes is performance auditing built on the proposition that a program should include activities or processes specified by one or more program stakeholder groups; these activities/processes may or may not result in desirable outcomes. (Post)positivism in performance auditing builds on the proposition that one reality does exist related to performance and that: (1) multiple accounts/perspectives of this reality are data that should be collected, including quasi- and non-experimental
Fig. 1. Classifying Theories of Program Evaluation: Focus on What is Explained and How Reality Regarding Evaluation is Viewed.

View of Reality and Research Methods: (Post)positivism

- Kotler
- Senge
- Campbell

Valuing activities/processes

- Stake
- Wholey

Valuing results

View of Reality and Research Methods: Relativism

Post Positivism and Focusing on Valuing Goal-Free Effects: Scriven

First, the body of work by Scriven (1967, 1974, 1986, 1995) focuses both on valuing effects resulting from the program implemented as well as judging the organization/processes performed. He defines evaluation as the science of valuing: “Bad is bad and good is good, and it is the job of evaluators to decide which is which” (Scriven, 1986). Scriven believes the evaluator’s ultimate task is to try very hard to condense the mass of data collected into one work: good or bad. “Sometimes this really is impossible, but all too often the failure to do so is simply a cop-out disguised as, or rationalized as, objectivity” (Scriven, 1971, p. 53). Scriven coined the identical concepts “goal-free” and “needs-referenced” evaluations. “The
evaluator’s job,” according to Scriven, “is to locate any and all program effects, intended or not, that might help solve social problems. Goals are a poor source of such effects. Goals are often vaguely worded to muster political support, and rarely reflect side effects that are difficult to predict” (quoted in Shadish et al., 1991, p. 80).

Scriven is a post positivist. He rejects positivistic ideas that reality is directly perceived without mediating theories and without perceptual distortion, that scientific constructs should be operationalized directly in observables, and that empirical facts are the sole arbiter of valid scientific knowledge (Shadish et al., 1991). Scriven argues that reality is independent of the observer, and it is possible to describe reality objectively. He does encourage the use of positivist research methods (e.g. experiments with test and control groups) to measure program impact; however, he stresses that practical constraints often limit the use of positivist research methods.

Note that in Fig. 1, Scriven is plotted right of center, toward valuing results. Scriven (1974, p. 16) himself provides the reason, “One way or another, it must be shown that the effects reported could not reasonably be attributed to something other than the treatment or product. No way of doing this compares well with the fully controlled experiment, and ingenuity can expand its use into most situations.” Scriven also believes in valuing activities/processes: he has developed an eighteen-point “Key Evaluation Checklist” that includes measuring functions and processes done (see Scriven, 1980, pp. 113–116).

Scriven also advocates the use of multiple methods and observers to reach convergence for learning facts about objective reality. He acknowledges that no single scientist ever observes reality completely and undistorted; everyone observes through biased perceptual filters. Consequently, according to Scriven, using multiple perspectives and methods helps construct a more complete and accurate view of objective reality than can be accomplished by applying any one research method. He calls this view “perspectivism,” and differentiates the view from relativism: “Perspectivism accommodates the need for multiple accounts of reality as perceptive from which we build up a true [objective] picture, not as a set of true pictures of different and inconsistent realities” (Scriven, 1983, p. 255). The original concept of “triangulation” (Denzin, 1984) is applied perspectivism: the use of multiple methods and observers to validate information about objective reality.

An important point to note here is that Denzin and other researchers (Denzin & Lincoln, 1994) have adopted a new definition for triangulation: the research process of using multiple observers and methods for assuring multiple viewpoints of reality (also see Stake et al., 1997). The proposition that an objective reality exists is not implied as part of this recent view of triangulation; this later view represents
a shift toward viewing reality from a relativist perspective. To avoid confusion, we refer to triangulation as applied perspectivism as “triangulation for convergence” and triangulation for gaining access of multiple viewpoints as “triangulation for divergence.”

**Positivism: True and Quasi-experimentation for Valuing Results**

Positivism focuses on achieving unbiased valuation of results of an implemented program. Positivism includes the assumptions: (1) that an objective reality exists; and (2) that we can measure accurately the effects of changing an independent variable (such as: expenditures levels of a marketing program or alternative advertising campaigns) on the levels of dependent variables (such as the number of inquiries, visits, length of stay, and total visitor party expenditures) can be accurately measured.

First fully developed by Fisher (1949), the purest form of positivism is the application of the “true experiment,” that is, the use of treatment groups and control groups with random assignment of subjects among the groups to achieve “internal validity,” a term coined by Campbell. “Random assignment promotes internal validity because it falsifies many competing interpretations – threats to internal validity – that could plausibly have caused an observed relationship even if the treatment had never taken place” (Shadish et al., 1991, p. 126, italics in original). Campbell and Stanley (1963) identify nine threats to internal validity in a true experiment. “Their explication is one of Campbell’s major contributions to social science and to evaluation” (Shadish et al., 1991, p. 126).

Quasi-experiments (Campbell, 1969) are novel research designs that rule out threats to internal validity when true experiments are not used. Quasi-experiments often include measuring the occurrence of a target variable (such as requests for a catalog, visits, length of stay, expenditures) when the program implemented to cause the target variable is absent. Thus, many quasi-experimental designs fill the need to measure impact by measuring a target variable when a program is present and absent. Campbell’s (1969) “Reforms as Experiments” is likely the most well known quasi-experiment.

Contrary to popular belief, true experiments to measure the impact of a program are used often for measuring the effect (i.e. impact) of marketing and advertising programs (see Banks, 1965; Caples, 1974; Woodside, 1996). Exhibit 1 presents an example of a true experiment applied to examine the effectiveness of a government tourism-marketing program. The data summarized in Exhibit 1 are fictitious because no such study exists in the literature on tourism advertising.
Exhibit 1. Positivism: Example of a Basic Design of a True Experiment with Simulated Data for a Government Tourism-Marketing Program. Note: R = random assignment of the 2,000 households to Group 1 or Group 2. A high inquiry rate occurs for Group 1: 100/1,000 = 10%, a lower inquiry rate occurs for Group 2: 10.1 = 1%. The “conversion rate” for Group 1 is 50/100 = 50%, the conversion rate for Group 2 is 5/10 = 50%. The average total number of nights for Group 1 is 40% higher than the average number of nights for Group 2 visitors ((3.5 − 2.5)/2.5) = 0.40. The average total expenditures for visitors from Group 1 is $300 greater than the average total expenditures for visitors from Group 2 ($750 − $450 = $300).

Almost all government tourism-advertisements include an offer for free literature (e.g. a “visitor’s information guide [VIG]”). The ads are created in part to gain inquiries, that is, requests, for this free literature. A VIG is an example of “linkage advertising.” Such linkage-advertising offers provide a mechanism for advertisers to identify “qualified prospects,” that is, people displaying an interest in buying the advertised product. Marketers sometimes use lists of such qualified prospects for additional unique marketing programs to convert qualified prospects into product buyers (see Rapp & Collins, 1994).

Note that the test displayed in Exhibit 1 is a comparison of five impacts of implementing the program vs. not implementing the program, for example, exposing target customers to advertisements to visit Texas vs. not exposing target customers to these advertisements. Because the two groups were created using random assignment with 1,000 persons per group, the two groups are identical statistically before the two treatment levels (i.e. program exposure vs. no program exposure) are administered. Adding more groups makes other comparisons possible, for example, between Program A vs. B vs. no program conditions. Such experiments are often referred to as “split-run” tests because a sample from a known population
(e.g. a sample of a list of newspaper, magazine, or cable TV subscribers) is split randomly to create two or more equal groups; then each group is exposed to a different treatment (see Schoenbachler et al., 1995).

Given the results in Exhibit 1, was the program effective? That is, did the program increase the levels of the five dependent variables that would not occur without exposure to the program? To answer this question with respect to a government tourism-advertising campaign (the largest budget item for most state tourism-marketing programs), first note that the inquiry rate is much higher for Group 1 vs. Group 2 (10% vs. 1%). We conclude that the campaign was effective in increasing the inquiry rate. Second, note that the “conversion rate” of visitors from inquiries was the same for both groups: 50/100 = 5/10. Thus, the conversion rate did not increase for Group 1 vs. Group 2.

Third, note that the rate of total visitors did increase for Group 1 vs. Group 2: 80/1,000 > 35/1,000. The difference in visitor rates is 4.5% (8 – 3.5 = 4.5). The rate of visits among households exposed to the program was more than 100% above the visitor rate of households not exposed to the program. Fourth, note that the average number of nights per stay is much higher for Group 1 vs. Group 2: 3.5 > 2.5. Also, the average total expenditure is much higher for visitors from Group 1 vs. Group 2: $750 > $450.

Hundreds of such true experiments have been done in industries other than government tourism-marketing programs (see Caples, 1974), surprisingly none have been reported in the travel and tourism literature (Woodside, 1990). In relation to testing the impacts of advertising campaigns, the costs of such experiments are relatively small – most likely less than 3% of the total advertising budget.

Rather than using true experiments or well-designed quasi-experiments, the evaluation of the impacts of government tourism-advertising programs has used one-group conversion studies almost exclusively. These conversion studies use no non-exposed, randomly assigned group. Thus, such studies cannot answer the basic issue: did the advertising program increase the number of visitors and dollar expenditures vs. what would have occurred without the advertising campaign? Such one-group conversion studies are useful for comparing the conversion rates and average dollar expenditures of different groups exposed to different government tourism-advertising campaigns (e.g. see Woodside & Motes, 1981) and competing advertising media (e.g. see Woodside & Soni, 1990), but such studies do not address the more basic issue of did the advertising program cause results that would not otherwise have been experienced.

Another major problem with conversion studies sponsored to date by government tourism-marketing agencies is the failure to measure the impact of advertising on households exposed to the advertising who do not inquire – shown
as variable \( c \) in Exhibit 1. Although the hypothetical example in Exhibit 1 indicates no increase in the visiting rate among non-inquirers exposed vs. not exposed to the advertising, a positive (and in some studies, a negative) impact is likely to occur.

Even though highly critical of both the activities and results of marketing programs, none of the seven performance-audit reports include mentioning the lack of high-quality research designs to measure advertising impacts. These performance audits do not include indications of knowledge of the core assumptions of positivism as a theory of program evaluation.

**Shift Away from Positivism toward Valuing Activities and Relativism Methods: Weiss**

Note in Fig. 1 that Weiss appears in two positions. “Weiss (1972) initially tried to use traditional experimental methods in evaluation, but the political and organizational problems she encountered in doing so sensitized her to the need to consider such problems more realistically” (Shadish et al., 1991, p. 183). In her later work on evaluation theory, she perceived evaluation as a political activity in a political context; she decided that evaluation is more useful if it suggests feasible action or challenges current policy.

I would like to see evaluation research devote a much larger share of its energies to tracing the life course of a program: the structure set up for its implementation, the motivations and attitudes of its staff, the recruitment of participants, the delivery of services, and the ways in which services and schedules and expectations change over time, the response of participants and their views of the meaning of the program in their lives (Weiss, 1987, p. 45).

Note that, unlike Scriven and Campbell, Weiss is more of a descriptive theorist than a prescriptive theorist of valuing. She advocates evaluations that report how stakeholders, such as program administrators, value their activities to assess whether or not changes are warranted/possible in policy decisions. Weiss’ primary goal for evaluations is the enlightenment of stakeholder groups – in thinking about issues, defining problems, and gaining new ideas and perspectives.

Scriven and Campbell focus on studying [valuing] the effects of possible solutions to problems; Weiss makes a role for studying all activities associated with social problem solving. . . . Second, Weiss has a more realistic view of how social programs work and change. Scriven and Campbell wrongly assume that programs will be significantly responsive to rational information, that ineffective programs will be replaced and effective ones retained, and that programs can be treated as uniform wholes implemented similarly across sites and times. Weiss explicitly rejects such assumptions (Shadish et al., 1991, pp. 216–217).
Partnering with Program Management to Achieve Performance Goals: Wholey

Note that Wholey is located in the bottom right quadrant of Fig. 1. While in principal endorsing the view that an objective reality exists and results can be measured using positivistic tools, Wholey (1977, p. 51) observes, “Practically, however, we must settle for substantially less in virtually all federal programs. An assumption is considered ‘testable’ if there exists test comparisons that the manager/intended user would consider adequate indication that observed effects were attributable to program activities.”

Wholey’s theory of program evaluation includes working closely with program managers. He writes, “Evaluators and other analysts should place priority on management-oriented evaluation activities designed to facilitate achievement of demonstrable improvements in government management, performance, and results” (Wholey, 1983, p. 30). Wholey and his colleagues advocate the reporting by the evaluator to the program managers about progress and results at all steps in the evaluation. “Final reports and exit briefings ought not be the places where significant findings are revealed for the first time. Evaluators should continually be sharing their insights, findings, and conclusions with staff, managers, and policy makers” (Bellavita, Wholey & Abramson, 1986, p. 291).

Wholey prefers evaluators to limit their role as program critics and to be team players working with management to improve programs. “Although evaluators clearly should not be asked to be cheerleaders for government, there is a legitimate point that they can play a more positive role in both shaping public opinion about government and helping agency managers run their programs” (Bellavita et al., 1986, p. 286).

From our examining the seven audits described below, government auditors completing performance audits of their governments’ tourism-marketing programs follow Wholey’s assumptions regarding valuing results using relativistic methods. However, the work and reports by the auditors do not follow Wholey’s views on partnering with program management. The state auditors in the U.S. did not follow Wholey’s recommendation of providing on-going reports of findings during the audit processes – the tone of the reports are confrontational, an emotional state that Wholey avoids in his auditing work.

Evaluating Activities and Using Relativistic Methods: Stake

“Stake advocates that case study methodologies be used to improve local practice” (Shadish et al., 1991, p. 271). Stake’s theory of program evaluation is closer to Weiss’s enlightenment approach than to Wholey’s partnering views. Unlike
Wholey, Stake does not emphasize concerns by legislators, although he describes such concerns in his evaluations.

Stake’s empirical work concentrates on evaluation of education programs. His theory of program evaluation includes the beliefs that evaluations should increase practitioner understanding and allow the practitioners (i.e. program managers) to solve any problems uncovered by the evaluation. Major concerns for Stake include: (1) how can evaluations be made to facilitate their use; and (2) how can evaluation reports be made credible? He argues that credible reports are more likely to be used than reports believed to be false. He argues that case studies are the best way to produce credible reports (see Stake & Easley, 1978; Stake et al., 1997).

Consequently, Stake is a strong advocate of the use of case-study methods. These methods include the use of interviews, observation, examination of documents and records, unobtrusive measures, and investigative journalism, resulting in a case report that is complex, holistic, and involves many variables not easily unconfounded. These written cases are informal, narrative, with verbatim quotations, illustrations, allusions, and metaphors.

Stake distinguishes between a preordinate evaluation and a responsive evaluation. “Many evaluation plans are more ‘preordinate,’ emphasizing: (1) statement of goals, use of objective tests; (2) standards held by program personnel; and (3) research-type reports” (Stake, 1980, p. 76). He advocates the use of “responsive evaluation” in place of preordinate evaluation. A program evaluation is responsive evaluation: “(1) if it orients directly to program activities rather than to program intents; (2) if it responds to audience requirements for information; and (3) if the different value-perspectives of the people at hand are referred to in reporting the success and failure of the program” (Stake, 1980, p. 77).

Stake has moved away from attempting to find objective reality and the use of triangulation of methods to converge on a view of such reality. Earlier he wrote that, “One of the primary ways of increasing validity is by triangulation . . . trying to arrive at the same meaning by at least three independent approaches” (Stake & Easley, 1978, p. C.27). As we have seen earlier, he now views the use of triangulation as a means of learning the multiple perspectives of reality. He states that there are only the realities that we construct: valid knowledge is that which bests portrays the richness and individual variations of each person’s created reality; and the case method is best suited to such portrayals.

Our meta-evaluation of the seven performance audits of tourism-marketing programs indicates that they do not match closely with the core propositions of Stake’s theory of program evaluation: these audits do not include rich, detailed, case studies of the programs examined; they do not include a relativist perspective of reality. Although triangulation methods are the primary research methods used
in these seven audits, the auditors use triangulation to achieve convergence rather than divergence.

**Objective Reality and Valuing Activities: Kotler**

Kotler’s (see Kotler, 1997; Kotler, Gregor & Rodgers, 1977, 1989) work on marketing audits may be viewed accurately as a post positivist theory of program evaluation similar to Scriven’s (1983, 1995) contributions: both Kotler and Scriven imply an objective reality exists and can be best learned using multiple research methods (i.e. triangulation to achieve convergence). However, Kotler focuses much more than Scriven on valuing activities rather than results. Thus, note the placement of Kotler in the upper left quadrant of Fig. 1.

*Kotler (1997, p. 777) defines a marketing audit to mean “a comprehensive, systematic, independent, and periodic examination of a company’s – or business units’ [or, policy program such as government tourism-marketing programs] – marketing environment, objectives, strategies, and activities with a view to determining problem areas and opportunities and recommending a plan of action to improve the company’s marketing performance.” He reports, “...most experts agree that self-audits lack objectivity and independence ... the best audits are likely to come from outside consultants who have the necessary objectivity, broad experience in a number of industries, some familiarity with the industry being audited, and the undivided time and attention to give to the audit” (Kotler, 1997, p. 777).*  

He continues, “The cardinal rule in marketing auditing is: Don’t rely solely on the company’s managers for data and opinion. Customers, dealers [i.e. firms carrying manufacturers’ product lines], and other outside groups must also be interviewed. Many companies do not really know how their customers and dealers see them, nor do they fully understand customer needs and value judgments” (Kotler, 1997, p. 779). Similar to Scriven’s “Key Evaluation Checklist,” Kotler provides checklists, rating forms, and open-ended questionnaires useful for the marketing audit he describes (see Kotler, 1997, pp. 776–782). Kotler’s “Components of a Marketing Audit” open-ended questionnaire focuses on evaluating the behavior of markets, customers [or clients], competitors, distribution and dealers, facilitators and marketing firms, and the public. Kotler’s audit form includes questions relating to valuing how well the executives of the audited organization are scanning the organization’s stakeholders (e.g. markets, industry, customers and competitors), planning decisions, implementing these decisions, and assessing profitability and costs.

In the following sections, we note specific areas of inactivity in most of the seven performance audit reports on how the government tourism-marketing
Assessing Alternative Methods of Performance Evaluation

programs are being managed. However, in our meta-evaluation, none of the seven performance-audit reports indicated use of checklists, rating forms, or open-ended questionnaires for valuing the scanning, planning, or implementing activities of the programs evaluated. The seven audits exhibit little evidence of systematic performance auditing for valuing the full range of management functions, as proposed by Kotler (1997).

Systems Thinking and Systems Dynamics Modeling as a Theory of Program Evaluation: Senge

While Senge (1990) and other systems researchers (e.g. Hall, Aitchison & Kocay, 1994) do not refer to program evaluation, their work clearly relates to valuing activities and outcomes of programs as systems. Similar to Wholey (1983) Senge does not suggest program evaluation to find fault. But Senge goes further than Wholey, “In mastering systems thinking, we give up the assumption that there must be an individual, or individual agent, responsible. The feedback perspective suggests that everyone shares responsibility for problems generated by a system” (italics in original, Senge, 1990, p. 78). Senge recommends developing a useful micro world (i.e. simplified representative view of the events and flows of relationships) of actual systems; he views everything in life as a system. A systems-dynamic modeling is an advanced form of creating micro worlds (see Hall, 1983). The view that customer behavior influences the behavior of program managers and program manager behavior influences the behavior of customers is the circular linkage included in a systems theory of program evaluation. Developing long-term relationship marketing programs with customers, including an interactive database marketing strategy with multiple contacts over several years, is a systems view of program valuing.

Note that Senge is located in a positivistic, valuing activity position in Fig. 1. However, systems researchers never advocate one reality as optimal; they advocate modeling the activity relationships among variables in a system and measuring the impacts of inputs and outcomes for each variable over time – using objectively-viewed indicators for each variable. Hall (1983, 1994) provides useful examples of such systems evaluation work.

A systems-thinking approach to tourism-marketing program evaluation includes identifying the participants, their activities, and the relationships among events occurring in implementing the program. Core assumptions in such systems-modeling work include the following propositions:

(1) all events are both causes and outcomes of other events;
(2) several feedback loops among two or more participants and events occur in the implementation of the program;
participants in managing and operating a tourism-marketing program have incomplete mental models of how the program functions – the human mind can not comprehend clearly the substantial complexities and nuances in a system; (4) the longer, indirect, paths in systems often have a total impact opposite of the effect of shorter, more direct paths; and (5) the principal objective of performance auditing of the program is to create useful systems models of the program and to test improvements in program outcomes via simulating the running of the program.

Figure 2 depicts the participants and main relationships (shown by arrows) for a systems theory of program performance auditing. Note that the main functions/activities performed by each of the eleven participant groups shown in Fig. 2 are identified. Note that the Department of Audits interacts chiefly with two principal parties: the Legislative Branch and the Department of Tourist Development. As Fig. 2 shows, audit researchers are likely to interact with other system principals, for example, travel industry firms.

The arrows in Fig. 2 indicate influence or scanning for information. Most relationships between principals in Fig. 2 indicate two-way influences. A few
Assessing Alternative Methods of Performance Evaluation

relationships in Fig. 2 are one direction to indicate information-scanning activities by a system participant. For example, under “Additional Third Parties: research firms,” customer data are collected; the research firm analyzes these data and prepares reports for the clients of the research firm. As depicted in Fig. 2, these clients may include the Department of Tourist Development or the Advertising Agency, or both these organizations.

A program-evaluation case study by Hall and Menzies (1983) illustrates the detail in systems performance auditing. Hall and Menzies (1983) describe a performance audit using micro-systems modeling of a leisure-sports organization – a curling club in Winnipeg, Canada. Details are described briefly here. Figure 3 summarizes the events (i.e. variables) and relationships for their systems model. The plus and minus signs on the arrows indicate the valence of influence between events. For example, increasing the annual fee (box 10) increases members’ dues (box 6) leading to increases in total profit (box 9). However, this short positive path (10→6→9) to increasing profit is misleading; increases in annual fees have strong, negative, and more indirect, links to profits. Please take a moment to identify all the indirect paths from annual fees (box 10) to total profit (box 9). (Hint: six of them occur starting with box 10 and going through box 7.)

Several simulations were run using the curling-club systems model for operating the leisure club for different decision scenarios. The simulation results included decreases in total profits following increases in annual fees, and increases in total profits with increases in bar and food revenue.

The propositions in a systems theory of program evaluation receive no specific attention among the seven performance audits reviewed in the meta-evaluation of government tourism-marketing programs. For example, none of the performance audits include evaluations concerning the lack of database marketing behavior for building long-term relationships with inquirers – inquirers responding to government offers of free travel information (e.g. visitor information guides). Because of the advances in the efficiency and effectiveness in database marketing programs and in system dynamics modeling (e.g. ALPHA/Sim, 1998), applications of systems thinking in performance auditing of government tourism-marketing programs are likely to occur – given that government performance auditors receive training in applications of database marketing and systems dynamics modeling.

Comprehensive Theories of Program Evaluation: Cronbach and Rossi

Two comprehensive theories of program evaluation are positioned in central locations in Fig. 1. These locations identify the contributions to program evaluation
Fig. 3. Policy Map of an Ice Sports Club Operation (Adapted from Hall & Menzies, 1983).
theory made by Lee Cronbach and Peter Rossi. Both Cronbach’s and Rossi’s views have much in common: they both advocate the use of multiple methods including positivistic and relativistic research methods. They both believe that no single paradigm (i.e. set of assumptions representing a mental model) for knowledge construction has sufficient empirical or theoretical support to dominate the field. Both want evaluators to study program description, causation, explanation, and generalization, sometimes including all of them in the same study, albeit with different priorities depending on circumstances (Shadish et al., 1991, p. 318).

The proposition that dependable knowledge is contingent on multiple conditionals captures Cronbach’s belief about objective reality. Rather than seeking a general estimate of objective reality, Cronbach seeks to reduce uncertainty and increase understanding of the complexities of relationships in given local contexts. Such complex, dependable, explanatory knowledge requires evaluators to place more weight on interpretive (relativistic), thick-descriptive methods rather than on methods traditionally espoused in social sciences – positivistic research methods, such as experimentation.

Cronbach and his colleagues (1980) make some highly perceptive observations about program administration and decision making in general in Toward Reform of Program Evaluation:

…rarely are decisions made rather than slipped into; rarely is there a single decision maker; rarely do data about optimal decisions take precedence over politics; rarely do all stakeholders give the project the same attention and meaning as evaluators; rarely is a report ready on time; rarely are decisions made on particular days; and rarely are evaluation results used to modify programs in ways that instrumentally link the modifications to the evaluation. Cronbach and his co-authors depict a world of politics and administration that undercuts the rational model and its image of clear command and optimal decision making (Shadish et al., 1991, p. 335).

This view of sensemaking matches well with the later work on sensemaking by Weick (1979, 1995). Weick views most decision making as following behavior rather than preceding behavior. He asks, “How can I know what I value until I see where I walk [what I’ve done]? People make sense of their actions, their walking, and their talking. If they are forced to walk the talk, this may heighten accountability, but it is also likely to heighten caution and inertia and reduce risk taking and innovation” (Weick, 1995, p. 183).

Consequently, Cronbach believes that evaluations must be crafted to the political system as it is rather than to an abstract model of how it should be. He suggests that it is unwise to focus on whether or not a project has “attained its goals” (Cronbach et al., 1980, p. 5). For Cronbach, evaluation is a pluralist enterprise. Evaluations should contribute to enlightened discussions of alternative paths for social action, clarifying important issues of concern to the policy-shaping community (Shadish
Rossi’s theory of program evaluation recognizes three kinds of studies by evaluators: (1) analysis related to the conceptualization and design of interventions (e.g., government social programs, such as tourism-marketing programs); (2) monitoring of program implementation; and (3) assessment of program utility (i.e., worth). Related to government tourism-marketing programs, the first kind of study would include asking the following question: what type of marketer-customer relationship should state government design and how should state government go about implementing such a design? Only one government audit, the Evaluation of the Australian Tourist Commission’s Marketing Impact 1993, comes close to raising this issue.

Related to the program monitoring—the second issue—Rossi believes that, “Program monitoring information is often as important or even more so than information on program impact.” He suggests that evaluators can help by monitoring “(1) whether or not the program is reaching the appropriate target population; and (2) whether or not the delivery of services is consistent with the program design specifications” (Rossi & Freeman, 1985, p. 139). Regarding program utility, Rossi and Freeman (1985, p. 40) state, “Unless programs have a demonstrable impact, it is hard to defend their implementation and continuation.” Going much further than other evaluation theorists, Rossi and Freeman devote a chapter to cost-benefit analysis.

Note in Fig. 1 that Rossi is positioned above the midpoint on the (post)positivistic-relativistic dimension. This positioning is based in part on his view that “Assessing impact in ways that are scientifically plausible and that yield relatively precise estimates of net effects requires data that are quantifiable and systematically and uniformly collected” (Rossi & Freeman, 1985, p. 224). Rossi’s theory of program impact measurement includes measuring effects in the absence of the program:

Adequate construction of a control group is usually central to impact assessment. Such groups should consist of multiple units comparable to the treatment group [group exposed to the tourism-marketing program] in composition, experiences, and dispositions. The preferred option is use of randomized controls, something that is often more feasible with innovations than established programs. When randomization is not feasible, other options are (from most to least preferred): matched controls, statistically compared controls . . . . Rossi also advises increasing the precision of impact evaluations by adding multiple pre- and post-measures and using large sample sizes (Shadish et al., 1991, p. 403).

Rossi also believes that assessing how well program innovations are implemented is crucial. “The components of the delivery system must be explicated and criteria of performance developed and measured” (Rossi & Freeman, 1985, p. 77).
“This [measuring] includes assessment of the target problem and population; service implementation; qualifications and competencies of the staff; mechanisms for recruiting targets; means for optimizing access to the intervention, including location and physical characteristics of service delivery sites [e.g. visitor information guides; websites; interstate visitor welcome Centers]; and referral and follow-up” (Chen & Rossi, 1983; quoted from Shadish et al., 1991, p. 401).

Conclusions on Theories of Program Evaluation Applied to Performance Auditing of Government Tourism-Marketing Programs

The seven performance-audit reports do not include mention of alternative theories or methods for program evaluation. They all report using a triangulation of methods, including interviews, document analysis, and site visits. Although not stated formally, the goal for using multiple methods for ensuring for divergent viewpoint data is not implied in these reports. Rather, the two objectives for using a triangulation of research methods in most of these reports include: (1) uncovering major problems in the actions and outcomes of operating the government tourism-marketing program; and (2) achieving convergence for presenting a Scriven-type summary view of the one, objective, reality of program activities and performance outcomes (i.e. “good” or “bad”).

For example, the “Audit Highlights,” the first page in the Tennessee 1995 performance audit of the Department of Tourist Development, addresses one activity found to be a problem by the auditors. “The Cost-Effectiveness of Using State Personnel Instead of Outside Contractors Should Be Analyzed” (all capitals in the original, Tennessee 1995). Similarly, the 1994–1995 performance audit of Missouri’s Department of Economic Development, Division of Tourism, reports, “Our review was made in accordance with applicable generally accepted government auditing standards and included such procedures as we considered necessary in the circumstances. In this regard, we reviewed the division’s revenues, expenditures, contracts, and other pertinent procedures and documents, and interviewed division personnel.”

The central theme of the Missouri audit report is on reporting improper activities and bad performance. The report includes written responses to these findings by the Missouri Division of Tourism. The North Carolina 1989 audit report of the Division of Travel and Tourism has the same theme: “In reviewing the activities of the division we found that the division had failed to answer 93,221 reader service inquiries during the first part of 1988” (North Carolina, 1989, p. 2). Note that this quote comes from the Executive Summary of the North Carolina State Auditor report. Although less adversarial, the 1985 Minnesota audit of tourism programs
of the Department of Energy and Economic Development (DEED) focuses on identifying shortcomings in activities of the program – including DEED’s lack of assessment of the return on investment of its tourism-marketing programs.

With the exception of the audit report for Australia’s tourism-marketing program, the performance-audit reports do not include a comprehensive coverage of the activities and results of government tourism-marketing programs. Again, with the exception of the Australian audit, the audits do not attempt to identify best practices in assessing performance of tourism-marketing programs. The audits do not describe other possible theories of program evaluation as foundations for performance auditing. The audits include only brief mentions to the theoretical grounding in accounting for the audits.

We conclude that the substantial majority of performance audits of government tourism-marketing programs are summarized accurately as post positivistic and are oriented towards valuing both activities and results. Related to the theories of program evaluation described earlier, these audit reports most closely follow the assumptions of Scriven’s theory of program evaluation. Figure 4 shows a summary of our meta-evaluations of the theories-in-use for the seven performance audits.

Notably, the only performance audit that was compiled jointly with an audit department and the department managing the tourism-marketing program is Australia’s performance audit compiled by “The Department of Finance, Department of Arts, Sport, the Environment, Tourism and Territories & the Australian

Fig. 4. Classifying Performance Audits of Government Tourism-Marketing Programs by Two Dimensions of Program Evaluation Theory.
Tourist Commission” (Australia, 1993, cover page). The executive summary of the Australian performance audit states, “This is the first comprehensive evaluation of the Australian Tourist Commission’s (ATC) marketing and promotional activities” (Australia, 1993, p. 1).

The performance-audit partnering of departments and the attempt to achieve a comprehensive program evaluation are examples of the evidence we used to locate the Australian audit among the theoretical assumptions of Wholey, Weiss’s later work, and Cronbach (see Fig. 4). However, several findings in the Australian audit report support the conclusion that the Australian auditors assume that an objective reality exists; the Australian report uses triangulation to achieve convergence, as well as divergence. The Australian audit report is the only one among the seven analyzed that includes references to tourism-marketing impact studies done by other governments. Thus, the theory of program evaluation covers Rossi’s assumptions and recommendations for estimating objective reality.

PROPOSITIONS FOR PERFORMANCE AUDITING AND META-EVALUATIONS OF PERFORMANCE AUDITS OF GOVERNMENT TOURISM-MARKETING PROGRAMS

Our review of the marketing and strategic management literature (e.g. Buzzell & Gale, 1987; Clifford & Cavanaugh, 1985; Huff, 1990; Kay, 1995; Kotler, 1997) and the reviewed theories of program evaluation leads us to the following propositions. We argue that the performance audits of government tourism-marketing programs that include the following traits are most likely to result in achieving the goals of government and public stakeholders:

- being comprehensive by including both (post)positivistic and relativistic research methods and both objective and subjective views of realities;
- valuing both program activities and results;
- carried out using a “partnering” (Wholey, 1983) approach by a “working party” (Australia, 1993) of representatives of three or five organizations assigned full-time to complete the audit.

The accounting paradigm for performance auditing as practiced now in the United States state audit reports is performing poorly in assisting in achieving:

- the program goals mandated by the state legislatures in creating: (a) the tourism-marketing departments; and (b) the performance auditing program by offices of legislative auditors;
knowledge of activities and processes performed in the government tourism-marketing department;
use of scientific methods for measuring the economic impact of government tourism-marketing programs;
the adoption and use of advanced database-marketing strategies in these tourism-marketing programs.

Accepting both objective and subjective realities is an example of applying Weick’s (1979) recommendations for sensemaking. “In environments where multiple contingencies arise, responses that are appropriate at one point in time may be detrimental at another. Cause mappings (see Huff, 1990) valid at one point in time may be invalid at another. Flexibility to deal with environmental changes is maintained if opposed responses are preserved” (Weick, 1979, p. 220).

Advanced database-marketing strategies include several characteristics, including the following features:

• learning the names, addresses and telephone numbers of customers and entering them into a database;
• having on-line computer files for each individual customer that can be accessed when a customer telephones or contacts the tourism department by email;
• based on knowing an individual customer’s leisure travel preferences and behavior, creating two-way communication strategies that include outbound contacts with customers beyond the one-time requests for brochures;
• and creating special member benefits for uniquely valuable visitors – such as frequent visitor programs.

Examples of such database marketing strategies exists in many travel-related industries, for example, car rental firms, credit card companies, and air lines.

Steps toward database marketing are indicated in a 1998 advertisement for Canada’s Atlantic Coast. (See Exhibit 2.) This ad’s design includes three specific features:

• four illustrations to create an image of the unique benefits offered by a trip to one or more of four Canadian Atlantic Coast provinces;
• an expensive free offer to gain inquiries for the “FREE 1998 Adventure Guide and Touring Planner. It’s over 100 full-colour pages of all the ways we make people feel good”;
• a tour package to gain direct sales for a touring visit including round-trip airfare, accommodation vouchers, compact car rental with insurance, and the “opportunity to extend your stay with extra-night rates.”
Assessing Alternative Methods of Performance Evaluation

**Illustrations:**
1. Beach, boulders and water at New Brunswick’s Bay of Fundy
2. Red-clay cliffs of Prince Edward Island
3. Waterfront buildings of Halifax, Nova Scotia
4. Map showing Canada’s Maritime Provinces

**Headline:**
CANADA’S ATLANTIC COAST

**Copy:**
Where adventure runs as high as the tides. Walk in the wake of the world’s highest tides or follow in the footsteps of the Vikings. Experience a seaside city filled with musical nights and Titanic sites. FREE 1998 Adventure Guide and Touring Planner. It’s over 100 full-colour pages of all the ways we make people feel good.

**FOOTLOOSE & FANCY FREE**
Here’s one of the many thrilling adventure packages you’ll discover on Canada’s Atlantic Coast: [a trip package including air fare from Boston to Saint John, New Brunswick; accommodations for 3 nights; compact-car rental for 3 days; toll-free reservation to “pre-book”; opportunity to extend stay with extra night rates]

**Direct Linkage Mechanisms:**
1. Toll-free telephone number: 1-800-565-2627
2. E-mail address: canadacoast@ns.sympatico.ca
3. World-wide website: www.canadacoast.com

Exhibit 2. Summary Description of Advertisement of Canada’s Atlantic Coast.

This ad to attract visitors to the Canadian Maritime Provinces includes three methods for inquiries by prospects: a toll-free telephone number; an email, and a website. Inquiries from this ad represent initial responses to the possibility of a highly interactive long-term relationship between the destination marketer and visitors. Several “outbound marketing” strategies can be used to reach inquirers, visitors, and repeat visitors for several years. A “frequency-marketing” program (e.g., frequent flyer memberships in the airline industry) could be created to stay in contact with frequent travelers and to promote trips highly profitable to the destination areas. Rapp and Collins (1988) provide an in-depth rationale for such strategies along with detailed descriptions of currently running database-marketing programs.

Create and implement database-marketing strategies because:

- a substantial share of tourist visits result in purchases to and around the destination area being promoted by governments;
- repeat-trip visitors represent an important target segment;
- advances in electronic technologies and software reduced the costs for operating such programs; and
many firms in many industries have completed the shift successfully to such “mass-customized” programs.

In particular, asking three questions is appropriate here:

- Given that most U.S. resident households will spend more than $1,000 for each leisure trip to Canada, are such database marketing strategies likely to be effective, efficient, and competitive?
- Do a substantial share of leisure visitors to Florida, California, Hawaii, Minnesota, Pennsylvania, Texas, or other states, provinces, and countries spend more than $1,000 on their trips?
- Which state and provincial governments are using advertising and inquiry responses to build database marketing programs with the described program characteristics?

The answers most accurate to these questions appear to be: yes, yes, and none. Only the answer to the third question may be surprising.

Figure 5 summarizes five specific activities that occur, or should occur, in a government tourism-marketing program. These activities include administering, scanning, planning, implementing, and assessing the performance quality of activities and program-system outputs. Including details on each of the five activities in a performance audit is a useful step toward achieving comprehensive evaluations of government tourism-marketing programs. None of the seven audit reports reviewed in the following sections reflects such a comprehensive review (the details leading to this conclusion are provided for each section reviewing the seven audits).

**Evidence of Scanning and Interpreting**

Scanning includes both formal and informal collecting and interpreting data on the behaviors, attitudes, and intentions of relevant stakeholders: clients (e.g.
Evidence of lack of formal scanning of stakeholder groups, especially potential customer markets, is an indicator of poor performance, arrogance, and the lack of management expertise. Daft and Weick (1984) note that patterns of scanning, interpreting, and learning vary across organizations as a function of their willingness to act in order to learn and their willingness to accept that the environment is difficult to analyze.

Related to the issue of formal scanning of relevant stakeholders, the issue of interpretation includes whether or not formal overall evaluation of the tourism-marketing program and “products” (i.e. state tourism services, benefits delivered, and visitor-perceived trip value) has been completed. In short, have program managers completed a formal overall evaluation of the program’s strengths, weaknesses, opportunities, and threats – a SWOT analysis? Is an annual SWOT analysis report prepared annually based on new scanning data?

**Evidence of Planning**

A formal, annual, marketing plan includes:

- an executive summary (usually prepared in 2–4 pages);
- a statement on the current marketing situation, including a SWOT and issue analysis, and the targeted segment of customers identified as potential visitors most likely to be influenced by the state’s tourism-marketing programs (4–6 pages);
- objectives that have been converted into measurable goals – to facilitate implementation and impact measurement (1 page);
- a marketing strategy – including a “what if analysis” and indicating the major decisions in the plan, for example, target markets and advertising strategy (e.g. a 200 page, fact-filled visitor’s information guide will be prepared and the offer for the free guide will be featured in all advertisements) (6–10 pages);
- planned actions – presents the details and schedules of what is to be done by whom, when, and how, and additional, related, special-events, marketing activities (4–6 pages);
- projected costs-benefits analysis – including the estimated expenditures by visitors in the state and state taxes collected from implementing the tourism-marketing program (2 pages);
• controls – indicating how the planned activities will be monitored to insure performance; including details of the methods used to measure the net impacts of the tourism-marketing program on inquiries, intentions, visits, revenues, and taxes (4 pages).

Does formal planning (i.e. a written statement explaining what you are going to do and why) really matter? If you believe the answer to be obviously yes, posing the question seems surprising. However, examining this issue is important – especially since none of the seven audit reports describes formal planning done by the tourism-marketing executives.

Armstrong (1982) conducted a meta-evaluation of all published research on evaluation of formal planning. In his evaluation, Armstrong asked, what is the value of formal planning? His analysis indicates:

Formal planning was superior in 10 of the 15 comparisons drawn from 12 studies, while informal planning was superior in only two comparisons. Although this research did not provide sufficient information on the use of various aspects of the planning process, stakeholders [in the formal planning process] provided mild support for having participation. Formal planning tended to be more useful where large changes were involved, but, beyond that, little information was available to suggest when formal planning is most valuable (Armstrong, 1985, p. 582).

The Usefulness of Written Reflection

We suggest including a “retrospective commentary” (in 2–4 pages) in the annual written marketing plan. This retrospective commentary can provide answers to the following issues: how much does what happened last year match with our strategic plans and action plans?

• What was planned that was executed well and poorly? Why?
• What was planned that did not get done? Why?
• What noteworthy events got done that were unplanned? Why?
• What did not get done but should have been planned and done? Why?

By including a retrospective commentary, program managers force themselves to reflect – to examine last year’s marketing plan and to ponder for a few moments about what happened and why (not necessarily, what things that went well and poorly, and why). Without a retrospective commentary, last year’s annual plan may never get noticed, or used, after it is written. Weick (1979, 1995) argues persuasively for reflection. “Reflection is perhaps the best stance for both researchers and practitioners to adopt if the topic of sensemaking is to advance” (Weick, 1995, p. 192). Weick (1995, pp. 24–26) further observes that, “. . . people can know what they are doing only after they have done it. . . . The creation of meaning is an attentional process, but it is attention to that which has already occurred. . . . Only
When a response occurs can a plausible stimulus then be defined [to explain its occurrence].”

The important point is that retrospective sensemaking is an activity in which many possible meanings may need to be synthesized, because many different projects are under way at the time reflection takes place (e.g., Boland, 1984). The problem is that there are too many meanings, not too few. The problem faced by the sensemaker is one of equivocality, not one of uncertainty. The problem is confusion, not ignorance. I emphasize this because those investigators who favor the metaphor of information processing (e.g., Haber, Ullman & Leifer, 1979) often view sensemaking, as they do most other problems, as a setting where people need more information. That is not what people need when they are overwhelmed by equivocality. Instead, they need values, priorities, and clarity about preferences to help them be clear about which project matters. Clarity on values clarifies what is important in elapsed experience, which finally gives some sense of what that elapsed experience means (Weick, 1995, pp. 27–28).

This meta-sensemaking view fits with the theory and empirical work in strategic management by Mintzberg (1978). Mintzberg defines strategy as observed patterns in past decisional behavior. Retrospection and research similar to Mintzberg’s studies on realized strategies lead naturally into mapping strategy (see Huff, 1990) and program evaluation using strategy mapping tools (e.g., Stubbart & Ramaprasad, 1988). Consequently, retrospection may be viewed usefully as a valuable major tool for both sensemaking and program evaluation theory and research.

Figure 6 is intended to be a useful sensemaking tool for retrospective analysis – in comparing planned strategies with realized strategies. Note that the bottom right-hand cell reflects the use of future perfect thinking applied to the immediate past: what did not happen that was unplanned but might have (should have) occurred?

In one study on using future perfect thinking for sensemaking, Boland (1984), gathered a group of film-lending executives in 1980, provided them with accounting reports prepared for 1982–1985, and asked them to imagine it was July 21, 1985, and then to discuss what the film service had become and why. Boland reported that a major outcome of the study was that in trying to understand what had been done in an imaginary future, participants discovered that they had an inadequate understanding of an actual past. The study uncovered disagreements about the nature and meaning of past events that people did not realize had impeded their current decision making (from Weick, 1995, p. 29).

Compared with not using such a sensemaking tool, answering the issues raised in all four cells in Fig. 6 promotes reflective thinking and likely will result in more useful, deeper, sensemaking. “All the mysteries become clearer [and some are solved] when we pay more attention to how people plumb the past and to what outcomes they have in hand when they do so” (italics added, Weick, 1995, p. 185).
If Mintzberg (1988) is correct in proposing that realized strategies rarely are the same as planned strategies, then both the program administrators and performance auditors need to attend to what actually gets done. Implementing includes accomplishing and not accomplishing the completion of events scheduled in the marketing plan. Implementation questions include:

- What was done?
- By whom?
- Where?
- When?
- How was it done (i.e. the detailed steps and with what observed levels of skills)?
- What occurred next (i.e. the event chain following the event)?

Marketing implementation has been defined to mean the process that turns marketing plans into action assignments and ensures that such assignments are executed in a manner that accomplishes the plan’s stated objectives (see Bonoma, 1985; Goetsch, 1993; Kotler, 1997). This definition is narrow and prescriptively
Assessing Alternative Methods of Performance Evaluation

oriented to effective marketing implementation. Our definition of marketing implementation focuses on what gets done – accomplishing and not accomplishing the completion of events scheduled and not scheduled in the marketing plan.

For purposes of sensemaking and performance auditing in program evaluation, three concerns about implementation are central: (1) learning the details done, and not done, in the streams of behavior that were actually implemented; (2) comparing these realized actions to the actions planned in the marketing plan; and (3) learning the immediate and downstream outcomes to the observed stream of implemented decisions and behaviors. The intended shift here is from prescription toward description.

For both program management and program evaluation, specific issues to be examined for describing implementation include the following points – implementation mapping, critical incidents, coaching and coordination, and nonevent analysis:

- Have service-performance-process maps (e.g. see Shostack, 1987), or causal policy maps (e.g. see Hall, Aitchison & Kocay, 1994), been created that accurately describe implemented behavior, including the downstream events occurring as a result of implementations? Such mapping tools help us to learn the intricacies and nuances that have occurred in recent implementations.
- What critical incidents (e.g. see Binter, Booms & Mohr, 1994; Price, Arnould & Tierney, 1995) occurred during the implementations of the program? What were the immediate and downstream influences (outputs) of these critical incidents? What combination of events (inputs) resulted in the critical incident?
- How many times did on-the-job coaching and coordination actions occur during program implementation? How many times did senior program executives and middle managers participate in on-line implementation of day-to-day activities in different departments of the program: weekly, monthly, quarterly, once per year?
- What unplanned events did not occur that should have been planned and implemented? Which of these unplanned non-events would have likely resulted in substantial favorable vs. unfavorable outputs for the program?

In summary, we emphasize three points. First, use ethnographic research tools to describe what is happening. Several worthwhile implementation research tools are available for creating “thick descriptions” (Geertz, 1973) of realized strategies. Several industries use such tools to achieve deep understanding – improved sense-making – of implemented strategies (see Arnould & Wallendorf, 1994). These tools are helpful for both program administrators and performance auditors.

Second, in reviewing the seven performance audits of the tourism-marketing programs, we find little evidence is found in using such tools. The reports do not
include detailed descriptions of the activities actually performed, or the policy maps of the decisions implemented.

Third, the argument that members of the legislative branch will not read such thick descriptions of implementation misses the more important point: the reason for implementation research is to improve sensemaking and decision making among program administrators and performance auditors. Meta-evaluation of implementation research helps to overcome Cook’s (1997, p. 48) impression: “My impression, perhaps mistaken, is that we evaluators are less deliberate and systematic than others in accumulating substantive findings and in drawing inferences about the programs and classes of programs that seem to work best.”

The U.S. findings of empirical studies on the reading and use of performance audit reports is consistent for both state and federal legislative members. The evidence is remarkably clear: these reports will not be read by legislators, with few exceptions (for a review of these studies, see Patton, 1997, especially Chapter 1). Even when read by legislators, the reports yield little to no impact. If the reports are going to be read and used to cause improvements in program performance, several stakeholders need to “buy in” to the findings, conclusions, and recommendations in the report. Possibly the only effective buy-in, performance auditing tool that works is Wholey’s (1983) partnering approach. The Australian (Australia, 1993) performance audit is an example of this partnering approach applied to tourism-marketing program evaluation.

Impact Assessment of the Performance Quality, Conformance Quality, and Quality-in-Use of Program Activities and Program System Outputs

We use the term, “performance quality” to refer to a program’s effectiveness and efficiency in delivering benefits to stakeholders, including clients (e.g. tourists and prospective tourists), public and tourist-industry related firms (in the form of improved services, wise use of tax-funded program expenditures, and increases in revenues and taxes paid by tourists), legislative and executive brands, and program managers and operators. Performance quality is distinct from conformance quality; conformance quality refers to product manufacturing and service delivery that meets mandated standards. Also, performance quality is distinct from “quality-in-use,” that is, customer beliefs about the quality of the services they experienced. See Garvin (1987) for a discussion on the multiple dimensions of quality.

All three dimensions of quality are useful for evaluating the impacts of a tourism-marketing program. Though not described in the seven audit reports reviewed in this meta-evaluation, quality-in-use data (i.e. customer beliefs about quality experienced) have been collected in government-sponsored evaluation research.
of tourism-marketing programs (e.g., see Dybka, 1987 for a Canadian study on U.S. residents’ quality-in-use beliefs about their visits to Canada). Conformance quality data are included in some of the seven audit reports; one example has been described briefly (in the North Carolina 1989 audit report: the non-delivery of the advertised visitor guide to 93,221 inquirers requesting the guide).

For achieving both triangulation for convergence and triangulation for divergence, using multiple indicators is best for measuring each of the three dimensions of quality. Indicators for measuring the effectiveness and efficiency of activities and systems in a tourism-marketing program are summarized in Exhibit 3.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Effectiveness</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Variable Measured)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Conformance measures: activities done in compliance to standards set</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2. Cost-benefit analyses, e.g., comparisons of cost per inquiry (CPI)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>3. Return on investment or net profit estimates, where net profit equals</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>4. Revenues for tourist-related industry firms generated due to tourism-</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>5. Revenue per inquiry (RPI) estimated from tourism-advertising conversion</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>6. Government revenues (taxes) generated from tourism-related industries</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>7. Tourist visits to the county, state, or country due to tourism-</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>8. Quality-in-use measure of tourists’ beliefs of quality of services</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>9. Ratio-of-ratio comparison, e.g., CPI/RPI?</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>10. Tourist awareness, attitude, and intentions to visit?</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

*Exhibit 3. Impact Assessment Indicators of Tourism-Marketing Programs.*
Exhibit 3 describes four indicators of efficiency and six indicators of effectiveness. Effectiveness refers to the performance quality of the activity, the quality-in-use of the program, and the value of revenues generated from the program (sales and taxes resulting from the government tourism-marketing program). Efficiency refers to conformance quality, profits, and ratio comparisons of results due to the program.

The first indicator in Exhibit 3 refers to conformance measures, as described earlier. The second measure is cost per inquiry (CPI). CPI is equal to the cost assigned to running a given advertisement in a given media vehicle in a given time period divided by the total inquiries generated by this ad placement. Also, CPI estimates are reported sometimes in comparing tourism ad placements in competing media vehicles (Woodside & Soni, 1990) and competing categories of media (e.g., see Woodside & Ronkainen, 1982). Estimates of the net profit to government resulting from tourism-marketing programs are available in the tourism program evaluation literature (e.g., Minnesota, 1985; Woodside & Motes, 1981; Chap. 4 in Woodside, 1996).

Tourists’ visits generated by the government tourism-marketing program (the seventh indicator in Exhibit 3) are estimated by research sponsored by the program managers for most programs. However, the performance audit report for Minnesota (1985, p. xii) notes that such research was not being linked to the return on investment (the third indicator) for Minnesota’s tourism advertising campaign.

Faulkner (1997) offers an additional indicator of tourism-marketing performance that is not included in Exhibit 3. Faulkner advocates computing two difference-score indexes:

1. “Market bias index,” that is, the share of a particular market compared to the share of a destination’s visitors overall; market bias equals zero if the two measures are equal;
2. an “index of change” in the visitors received from each market relative to the change in that market overall; this change index equals zero if the growth rate of visits from an origin is equal to the overall growth rate of trips made by that origin.

Faulkner (1997) displays the computed indexes for each origin in two-dimensional exhibits.

However, we do not recommend Faulkner’s approach for measuring tourism-marketing performance. Using indexes based on difference scores causes severe measurement problems; consequently, most psychometric experts recommend that difference scores not be used (see Teas, 1993). For a given origin, such indexes do not provide estimates of market share or rate of growth. Neither do such
indexes provide a measure of impact (i.e., effect size) of the tourism-marketing program. Thus, for Australia as a destination, a conclusion that Japan as a market bias index equal to 5.3 and an index of change equal to 0.6 provides little information and misleading implications. The impacts of presence vs. absence of specific marketing actions and expenditures on market share and change are not included in such index analyses.

Moreover, market share is not a particularly useful dependent or independent variable in performance measurement. “You’re better off if you create and develop [profitable market] niches” (Clifford & Cavanaugh, 1985, p. 17). When focusing on market share, judge the impact of changes in market share relative to a destination’s two key competitors – with and without the presence of specific market actions in a given tourism-marketing program. For example, the following steps can be used with market share as the dependent variable.

(1) Identify ten specific markets by a combination of origins and lifestyles.
(2) Estimate market shares for a destination for each of the ten markets for the three years prior to administering marketing actions to specifically attract visits from five of the ten markets.
(3) Randomly assign five of the ten markets to receive unique marketing programs to attract visitors from these markets to the destination; the other five markets represent “control” markets – they receive no unique marketing programs to stimulate visits.
(4) Continue to estimate market shares for each of the ten markets for the year that the marketing actions are implemented in the five randomly assigned markets.
(5) Continue measuring market share for fifth and sixth year even if the marketing actions are implemented only in year four.
(6) Using multiple regression analysis, estimate the effect size of the impact of the marketing actions on market share.

Figure 7 summarizes the results of a hypothetical example of following these five rules. Note in Fig. 7 that market share for destination X averages about 5% over years 1–3 for both the five treatment group markets and the five control group markets. The average market share for X is close to 11% during years 4–6 for the treatment group markets exposed to unique marketing actions in these markets. However, the average market share for X is still about 5% during years 4–6 for the five markets in the control group. An analysis of the findings in Fig. 7 for competitor Y confirms the view that the marketing actions performed by X during years 4–6 influenced market shares: Y’s market share declined systematically in years 4–6 compared to previous years in the treatment group markets but not the control group markets.
Fig. 7. Effect on Market Shares for Destination X and Competitor Y of the Marketing Actions by Destination X. Note: In this hypothetical example, systematic increases in market share for X are associated with the introduction of the marketing actions in years 4–6 in the treatment markets vs. years 1–3; in the markets in the control group no systematic change in market share for X occurred over the six years. Systematic market share decreases for competitor Y occurred in the treatment group but not the control group. Such findings represent strong evidence that the marketing actions implemented by X were effective in increasing market share for X and decreasing market share for Y.

Scientifically Measuring the Impact of Advertising and Marketing on Visits

The Minnesota (1985) performance audit of the state’s tourism-marketing program included more knowledge of positivistic research for valuing results compared to the other six performance audits. For example, the Minnesota audit identified severe shortcomings of return-on-investment findings in the budget requests made by the tourism-marketing program director; these shortcomings are described in the results section. The ROI findings, communicated by the program director, were based on the results of a conversion research study done: Following the tourism season in Minnesota, a questionnaire was mailed to inquirers asking if they had visited Minnesota. Two shortcomings noted in the audit report are worth emphasizing here:

The study assumed that all persons who said they had vacationed in Minnesota had come as a result of the office of tourism’s promotion campaign. As a result, the study overestimated the return on investment because many would have undoubtedly vacationed in Minnesota regardless of the tourism promotion campaign. Furthermore, even for those who did not decide until after receiving tourism brochures, it is difficult to measure how much influence the brochures and advertising had on their vacation planning. . . . the study did not attempt to measure how many people vacationed in Minnesota as a result of the state’s advertising campaign even though they
did not request information from the tourism office. In addition, the study did not attempt to measure future benefits due to repeat visits or word-of-mouth advertising (Minnesota, 1985, pp. 91–92).

Unfortunately, the Minnesota performance-audit report does not indicate knowledge of true and quasi experiments for separating the impact of advertising from other factors on visits and expenditures (the positivistic theory valuing results and the empirical studies of Banks, 1965; Campbell, 1965; Caples, 1974; Ramond, 1966). In fact, the Minnesota performance audit reinforces the inaccurate and harmful myth that “It is not possible to separate the effects of tourism promotion from many other factors which affect tourism or travel spending, such as changing vacation preferences and changing business vacation patterns” (p. 90). True experiments are designs used to make such separations scientifically, to separate the impact of a new drug from other influence factors and medical research uses such designs widely (e.g. Salk and Sabin polio vaccine drug tests in the 1950s, see Woodside, 1990). In addition, such true experiments are used widely for measuring the unique impact of advertising on purchases; Caples’ (1974) work is the seminal work on this topic.

Unfortunately, the use of true experiments has never been reported for measuring advertising effectiveness or efficiency by government tourism-marketing programs. The most likely reasons for not using scientific designs to measure impact include the following points.

1. Lack of training and knowledge of the program evaluation literature in general among both program managers and performance auditors, especially lack of knowledge of the literature on measuring advertising effectiveness.

2. The argument that the research costs are prohibitive – a false argument because true experiments can be built into the planned advertising campaign; moreover, the cost of a running a true experiment are no greater than the research costs of conversion research studies.

3. Comfort with conversion research methods and “aided recall” advertising effectiveness studies.

4. Fear and high-perceived risk of using another type of research study – especially if the results might indicate the advertising had little impact on causing visits that would not have occurred without the advertising.

By using a true experiment, program evaluators can scientifically estimate the separate impacts of exposures both to scheduled-media placed ads and to the visitor information guide. Such measurement includes estimating several influences: (1) being exposed vs. not being exposed to the advertising; (2) being exposed vs. not being exposed to a visitor’s information guide (VIG); (3) the interaction effect of
both being exposed and receiving a VIG vs. the additive effects of ad-exposure only or VIG-exposure only. The required true-experiment design to measure these impacts includes intentionally not exposing some identified prospects to the ads and the VIG, while exposing others, identified prospects to the ad only or the VIG only or both the ads and VIG. This research design is an expanded version of Exhibit 1. The expanded version is shown as Exhibit 4.

Because program and advertising managers may be unwilling intentionally not to fulfill requests from prospective visitors for free literature, two literature packets should be used: a minimal packet (a skinny brochure that is as close as allowable to a no information-treatment condition) and the regular packet (a thick, 50–200 pages, visitor’s information guide full of details on special events, accommodations, restaurants, shopping, historical places, local touring routes, etc.). Using the research design displayed in Exhibit 4 answers three central questions.

- How many visits (and revenues, taxes, profits, etc.) resulted from the prospect-inquirers receiving the regular thick VIG vs. receiving the skinny brochure?
- How many visits resulted from the advertising program among persons exposed to the advertisements that did not request or receive the free VIG?
- How many visits (and revenues, taxes, profits, etc.) occurred because of the advertising program that would not have occurred without the program?

Taxes paid resulting directly from visits of tourists are government tourism-related revenues. What are the total direct tax-revenues resulting from the tourism advertising program? The tourism advertising impact on tax revenues can be estimated from Exhibit 4. Total tax revenues for each group:

Group 1: 
(4 visits × $12) = $48
(50 visits × $9) = $450 
Total tax revenues: $498

Group 2: 
(20 visits × $24) = $480
(100 visits × $18) = $1,800 
Total tax revenues: $2,280

Group 3: 
(30 visits × $42) = $1,260
(200 visits × $18) = $3,600 
Total tax revenues: $4,860

Consequently, total tax revenues due to the regular advertising campaign vs. no advertising campaign equals $4,362 (4,860 − 498 = 4,362). The advertising campaign net tax revenues are $4.36 per 1,000 target customers exposed vs. not exposed to the advertising campaign (4362/1000 = 4.362). The combination of the ad plus the minimal brochure offer also increases tax revenues compared to the tax revenues resulting from no advertising campaign (2,280 − 498 = 1,782). The ad plus minimal brochure campaign results in a net tax revenue gain of $1.78
Exhibit 4. True Experiment Design to Measure the Effectiveness of Schedules-Media Advertising and Brochure Literature Sent to Prospects. Note: VIG = visitor’s information guide, e.g. 100 pages. Note that the “conversion rate” of inquirers into visitors for each of the three groups is greater than 29%: 4/10 = 40% for Group 1; 20/50 = 40% for Group 3; 30/100 = 30% for Group 3. The conversion rate is lower for non-inquirers for each group: 50/90 – 5.5% for Group 1; 10.5% for Group 2; 22% for Group 3. Did advertising cause inquiries that would not have been made without the advertising? Yes: the inquiry rate was 1.0% for Group 1; 5.3% for Group 2; and 10.0% for Group 3. Use of this design requires high response rates (i.e. >50%) to a questionnaire sent to all members of the 3 groups; the questionnaire should not bias respondents by identifying the state sponsoring the study. Information regarding visiting several different states should be collected in the study to increase interest in answering the survey; to hide identity of the state sponsoring study.
per 1,000 target customers exposed compared to not having a tourism advertising campaign.

Which of the two advertising campaigns should be used? The answer depends in part on the cost of using the regular VIG vs. the lower cost of using the minimal brochure. We recommend performing a return-on-investment (ROI) analysis of each campaign. What is the net tax profit to the state for each of the three groups? Make some reasonable assumptions concerning costs. For example, assume that the total cost for creating, handling, and postage for the VIG is $4.00 per mailing, the per unit total costs for the minimal brochure is $1.25, and the cost of advertising in the schedule-medium is $400 per 1,000 subscribers for each ad. Addressing the ROI question and related issues (e.g., the increase in number of local areas toured by visitors receiving a VIG vs. visitors not receiving a VIG) is possible when powerful research designs are used to measure the impacts of tourism-marketing programs. (See Woodside, 1996 for a detailed real-life example of such effectiveness and for efficiency estimates for the tourism-marketing campaign for Prince Edward Island.)

Note that economic multiplier effects would need to be utilized in estimating the total impacts on revenues and taxes of the advertising campaign. Because the purchases by tourists result in purchases by tourism-industry service providers, and thus, a ripple effect of purchases occurs, we suggest a multiplier of 2.0–2.5 for estimating the total impact on revenues.

An additional point is worth noting from the hypothetical results shown in Exhibit 4. The impacts of a thick (100+ pages), detailed, VIG include substantial increases in the:

- average length-of-stays of visitors;
- average total expenditures of visitors;
- average number of local regions visited;
- intentions to return;
- and satisfaction with the services provided and the quality of the visit (for empirical evidence supporting these conclusions, see Chapter 4 in Woodside, 1996).

**Evidence on Administering**

As shown in Fig. 5, evidence on administering refers to vision-values, mission, training, coaching, and coordinating. Evaluations of success and failure in business organizations conclude that these administering issues are important (e.g., see Clifford & Cavanaugh, 1985; Scott, 1998). In a large-scale U.S. study of successful midsize companies, the senior executives in the firms studied:
... set forth vividly the company’s guiding principles – defining the ways value is to be created for customers, the rights and responsibilities of employees, and, most important, an overall affirmation of ‘what we stand for. Philosophy and values are hard to measure – but their value can’t be overstated. A statement of beliefs alone will not, of course, make a successful enterprise. Credos are an articulation of culture – not a substitute for it (Clifford & Cavanaugh, 1985, p. 13).

Clifford and Cavanaugh do, however, provide some concepts for measuring vision and values:

- earned respect – a sense that the enterprise is special in what it stands for;
- evangelical zeal – honest enthusiasm for creating and implementing plans;
- habit of dealing people in – communicating just about everything with everybody;
- view of profit and wealth-creation as inevitable by-products of doing other things well;
- leadership – executives show extraordinary commitment to the business, often to the point of obsession.

Related to these concepts, and the issues of training, coaching, and coordinating, are specific actions by program executives for achieving high levels of occurrence and to learn how administering impacts the four variables shown in the middle of Fig. 1. The senior program executives might ask:

- Do we have a written statement of vision and values unique for our state’s tourism-marketing program?
- Do our employees, visitors, and legislators perceive us to have evangelical zeal?
- Do we communicate just about everything with everybody?
- What formal training programs have employees and executives completed this past year? What impacts result because of their participation in these training programs?
- Are we coordinating enough? Are we meeting weekly for coordination and for sensemaking? How often do executives communicate personally with prospective visitors and actual visitors?
- Have managers completed formal training on measuring program performance?
- Are managers and employees sometimes uncertain about what they really know about markets, prospective visitors, visitor behavior and attitudes? Or, do they think and act as if they have the answers and external data on markets and customers are unnecessary?

What activity results in improving coordination and sensemaking? Answer: participating – sharing and revising thoughts – in meetings. Sound sensemaking practice requires that people need to meet more often.
That implication arises from the reluctance with which people acknowledge that they face problems of ambiguity and equivocality, rather than problems of uncertainty . . . . What will help them is a setting where they can argue, using rich-data pulled from a variety of media, to construct fresh frameworks of action-outcome linkages that include their multiple interpretations. The variety of data needed to pull this difficult task are most available in variants of the face-to-face meeting (Weick, 1995, pp. 185–187).

Thus, increase the quality of performance audits by measuring how often team meetings are held in the various departments of the tourism-marketing program, and the decisions and actions that follow from such meetings.

Work of the Performance Audit Team

For a comprehensive performance-audit report, include evidence that the work of the program-evaluation team included data collection on and interpretation of all five issues summarized in Fig. 5. Figure 8 illustrates this recommendation.

Note that Fig. 8 includes a sixth topic area: “Scanning other States and Industries by the Audit Team to Learn Best Practices.” We are suggesting here that a comprehensive performance audit include such scans by the audit team. Only two of the seven audits below (i.e. Australia, 1993; Minnesota, 1985) included such scanning.

A comprehensive meta-evaluation would include evaluations on several comprehensive performance-audit reports – including describing the presence or absence of each of the six areas described. A comprehensive meta-evaluation would include commentaries on:

Fig. 8. Performance Audits of Marketing Department.
Assessing Alternative Methods of Performance Evaluation

Fig. 9. Meta-evaluation Research of Performance Audits of Marketing Departments.

- the theory-in-use as displayed in each performance audit report;
- the scanning done by the audit team of other states and markets;
- the comprehensiveness of the performance audit reports;
- the subsequent impact of each performance audit report.

Figure 9 summarizes such comprehensive work by meta-evaluation researchers. Note that the box at the top in the middle of Fig. 9 refers to scanning by the meta-evaluation research team of other meta-evaluations (e.g. the meta-evaluation by Armstrong [1982]).

Reviews of additional meta-evaluation research studies on issues in program evaluation are likely to result in additional insights and nuances for theory and practice of program evaluation. Damson’s review (reported in Patton, 1997) of the legislative use of 21 program evaluations is another example of scanning to uncover other meta-evaluations.

METHOD FOR ASSESSING PERFORMANCE AUDITS OF GOVERNMENT TOURISM-MARKETING DEPARTMENTS AND PROGRAMS

This section describes the development of two instruments for content analyzing the performance audit reports. This section also describes the procedural details of acquiring and analyzing the audit reports.
Content Analysis Measures

Based on the literature review of theories of program evaluation and propositions developed for evaluating performance audits of government tourism-marketing programs, we developed two separate content measurement forms for analyzing such performance audit reports. Content analysis is research techniques for the objective, systematic, and quantitative description of the manifest content of communication (see Berelson, 1952; Kassarjian, 1977 for descriptions of content-analysis studies).

Exhibit 5 shows the first content-measurement form, covering five categories of issues. The issues in Exhibit 5 are questions that are answered from reading and content analyzing each performance audit report.

Most of the specific issues included in Exhibit 5 have been described earlier. However, one concept in the “other meta-evaluations” section deserves elaboration: “meta-analysis.” Meta-analyses are tools and theory used to accumulate knowledge across studies. “Scientists have known for centuries that a single study will not resolve a major issue. Indeed, a small sample will not even resolve a minor issue.
Thus, the foundation of science is the cumulation [sic] of knowledge from the results of many studies” (Hunter & Schmidt, 1990, p. 13).

Meta-analyses often include numerical estimates of the “effect size,” the size of the impact, of an independent variable on a dependent variable. For example, one effect size would be the estimated share of inquirers responding to the advertised offer of a free VIG who converted into visitors who would not otherwise have visited. Given the large number of conversion studies in the scientific tourism literature, surprisingly no meta-analysis of reported conversion shares has been reported. Even though these studies do not attempt to measure the unique impact of advertising vs. no advertising inquiry-to-visitor conversions, they do report shares of conversion that can be analyzed using meta-analysis tools (e.g. see Chapter 7 in Hunter & Schmidt, 1990).

The second content analysis form includes multiple items for each activity/decision topic described earlier in the section on propositions for performance auditing tourism-marketing programs. These topics include: scanning; planning; implementing; program self-assessment of activities and impacts; and administering. Exhibit 6 includes the specific items included for the content analysis of each of these five topics.

Related to items in Exhibit 6, a core issue related to a meta-evaluation of performance audits is whether or not data are collected. One reason for not collecting such information is that the performance audit was planned to be limited in scope – what Wholey (1983) refers to as a “rapid feedback evaluation.”

However, our suggestion is that all performance audits of tourism-marketing programs can and should include some data collection, evaluation, and reporting for all five major activities summarized in Exhibit 6. Even with a performance audit deadline of less than two weeks, such data can be collected, analyzed, and reported – while the performance auditors still focus mainly on very specific issues for a rapid feedback evaluation. In particular, we offer the suggestion because performance audits of tourism-marketing programs are usually completed less than once per decade, and dramatic changes and improvements in tourism-marketing programs are unlikely to occur without deep sensemaking learned from formal, broad-coverage, performance audits.

**Procedure**

We conducted searches of U.S. state government websites to learn of the completion of performance audits of tourism-marketing programs. We found six reports. We contacted state audit offices and requested copies of each of the reports. In addition, in August 1993, Peter Valerio, Manager, Strategic Analysis Unit of the
<table>
<thead>
<tr>
<th>Activity/Decision</th>
<th>Level Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scanning:</strong></td>
<td>Did Auditor Collect?</td>
</tr>
<tr>
<td>Evidence of formal scanning report?</td>
<td></td>
</tr>
<tr>
<td>Formal SWOT assessment done?</td>
<td></td>
</tr>
<tr>
<td>Market-by-market analysis: S/D?</td>
<td></td>
</tr>
<tr>
<td>Trends identified formally?</td>
<td></td>
</tr>
<tr>
<td>In-depth research &amp; knowledge of customer segments/targets?</td>
<td></td>
</tr>
<tr>
<td><strong>Planning:</strong></td>
<td></td>
</tr>
<tr>
<td>Up-to-date written marketing plan for tourism-marketing program?</td>
<td></td>
</tr>
<tr>
<td>Marketing plan includes written milestones of schedule actions?</td>
<td></td>
</tr>
<tr>
<td>Evidence of “what if” reflection/analysis?</td>
<td></td>
</tr>
<tr>
<td>Formal testing of alternatives? Before and/or after execution?</td>
<td></td>
</tr>
<tr>
<td>Formal projections of cost/benefit analysis of planned actions?</td>
<td></td>
</tr>
<tr>
<td><strong>Implementing:</strong></td>
<td></td>
</tr>
<tr>
<td>Evidence of on-the-job coaching and coordination/feedback?</td>
<td></td>
</tr>
<tr>
<td>Have service-performance process maps been created/used?</td>
<td></td>
</tr>
<tr>
<td>Written monitoring guidelines established/followed?</td>
<td></td>
</tr>
<tr>
<td>Vital competencies done in-house?</td>
<td></td>
</tr>
<tr>
<td>Formal system for tracking/responding to critical incidents?</td>
<td></td>
</tr>
<tr>
<td><strong>Activities/Impact Assessing:</strong></td>
<td></td>
</tr>
<tr>
<td>Realized/planed strategy match?</td>
<td></td>
</tr>
<tr>
<td>Multiple dependent impact measures?</td>
<td></td>
</tr>
<tr>
<td>Cost/benefit analysis?</td>
<td></td>
</tr>
<tr>
<td>True experiment of marketing impact?</td>
<td></td>
</tr>
<tr>
<td><strong>Administrative:</strong></td>
<td></td>
</tr>
<tr>
<td>Written mission statement prepared?</td>
<td></td>
</tr>
<tr>
<td>Annual quantitative objectives established?</td>
<td></td>
</tr>
<tr>
<td>Evidence of training, coordination, coaching by senior executives?</td>
<td></td>
</tr>
</tbody>
</table>

*Exhibit 6. Key Issues in a Meta-evaluation of Performance Audits of State Tourism Development Departments. Note: SWOT = strengths, weaknesses, opportunity, threats; S/D = supply and demand analysis; Levels: L1 = found to be poorly done; L2 = found to be done.*
The seven performance audit reports were evaluated using two content-analysis instruments. Two researchers, trained in theories and research methods of program evaluation, prepared written answers in response to the written questions in the two instruments. Two researchers worked separately in reading the reports and writing their answers. They then compared their answers and resolved conflicts through discussions and referring to the reports until they reached consensus. Complete agreement among the written answers was ranged 77–92% across the ten topics in Exhibits 5 and 6.

Note that the items included in “Other Meta-evaluations” in Exhibit 5 represent a self-assessment among the meta-evaluation researchers; not surprisingly, agreement was 92% for the written answers to these items. Part of the concern with self-assessments is the limited range of observed findings. This concern can be addressed by requesting additional researchers, knowledgeable about program evaluation theory but not involved directly in the current study, to serve as external judges. Each of these external judges would read some (or all) of the performance audits, review the completed content analysis instruments, and provide comments on the accuracy of the completed forms. This external judging procedure was not used in the study reported here, and not including such an external review step represents a limitation of the study.

**A META-EVALUATION OF A LIMITED FOCUS AUDIT**

*The 1996 Missouri Performance Audit Report of the Division of Tourism*

We begin with an evaluation of an audit report limited in focus. Margaret Kelly, CPA, State Auditor, signed the letter of transmittal for this report. The contents of the letter included the following statements. “We have conducted a special review of the Department of Economic Development, Division of Tourism… As part of our review, we assessed the divisions’ management controls to the extent we determined necessary to evaluate the specific matters described above and not to provide assurance on those controls.”

The report itself includes the following sections:

- history and organization of the Division of Tourism (3 pages). This section includes a statement of purpose: “The purpose of the division is to promote the state’s travel industry by encouraging visits by out-of-state vacationers and by encouraging Missourians to vacation in their home state”;
• a summary of findings (1 page);
• a “presidential library funding” section (3 pages); summary: $1 million grant made improperly by the division; “... state lost potential interest revenue by disbursing the grant monies before expenditures were incurred...” The not-for-profit corporation receiving the $1 million from the Division invested more than $800,000 of these funds into fixed income investments;
• a “sales tax funding” section (3 pages); summary: funding to the Tourism Supplemental Revenue Fund exceeded the statutory limit by $4 million;
• an “expenditures and contracts” section (5 pages); summary: several different payments irregularities found; bid documentation insufficient;
• a “sponsorships” section (2 pages); summary: “Written contracts were not obtained to specify what the division would receive in exchange for paying over $33,000 for promotional items and sponsorship fees to promote tourism at various events.”

Each section included a “WE RECOMMEND” subsection. For example, for the first section, “WE RECOMMEND the division seek reimbursement of the unexpended portion of the [$1 million] grant.” Each of these sections included specific findings and recommendations by the auditor. Also, “AUDITEE’S RESPONSE” subsections were included for each of the sections. The first of these responses begins by implying that the Division of Tourism did not commit an error in paying the $1 million: “The Division of Tourism clearly received $1 million for the purpose of funding the Truman Presidential Library in Independence, Missouri and made the payment to the library in compliance with the state of Missouri procedures.” However, the next sentence in the division’s response indirectly addresses the auditor’s concerns about making grants, “In the future, the Missouri Division of Tourism will ensure that the contracts are written for all significant expenditures and that disbursements as a result of these contracts are made only as expenditures are incurred.” Nothing is stated in the division’s response to the auditor’s recommendation for the “division to seek reimbursement of the unexpended portion of the grant.”

Claims and counterclaims are made by the auditor and auditee in the second section, “Sales Tax Funding” with both quoting portions of state regulations. For example, the “AUDITOR’S COMMENT” following the “AUDITEE’S RESPONSE” is the following statement:

While Section 620.467.2 RSMo 1994, states “[H]owever, such transfer in any fiscal year other than fiscal year 1996 shall not exceed three million dollars.” Therefore, not only is there no authority to transfer anything in excess of $3 million to the Tourism Supplemental Revenue Fund, in any fiscal year other than 1996, that is expressly prohibited by statute (Missouri, 1996, pp. 12–13).
The next section of the 1996 Missouri performance audit is entitled: “FOLLOW-UP ON STATE AUDITOR’S PRIOR RECOMMENDATIONS.” This section reviews several control problems noted in the 1994 audit report and the actions taken, if any, by the Division of Tourism between 1994 and 1996.

For example, the 1994 report includes the following findings. (1) Overpayment “by at least $50,876 to the advertising agency and the division made various payments to the advertising agency before the payments [made by the advertising agency to media firms] were actually done.” (2) “The Department of Corrections (DOC) provided telemarketing services to the division. The division did not request copies of the actual telephone bills paid by the DOC to determine the validity of the amount reimbursed to the DOC by the division.”

The auditee’s responded to these criticisms by stating that the “division received a $20,823 credit from the advertising agency.” The remaining amount of overpayments found by the auditor “was attributable to personnel classification errors” (p. 22). Also, the division discontinued the telemarketing contract with the DOC.

The final section of the audit report include three appendices: a schedule of appropriations; a comparative schedule of expenditures for five years (1992 through 1996); and a “comparative schedule of disbursements, transfers and changes in cash and cash equivalents” for 1995 and 1996. Appendix 2 shows advertising expenditures as 76% of the 1995 expenditures: ($5,950,181)/($7,854,393) = 75.76%. For 1996, if the $1 million library grant is excluded as non-recurring special funding, the advertising to total-expenditures ratio is 70%.

Meta-evaluation of the 1996 Missouri Performance Audit Report

The 1996 performance audit of the Division of Tourism by the Offices of the State Auditor of Missouri is concisely written with specific recommendations included. The report is exceptional in describing the history, organizational structure, several annual budgets, and expenditures. Moreover, the report is unique in providing a follow-up report on the findings, recommendations, and actions taken and not taken in a related previous performance audit report. Carrying out two performance audits of a state agency in less than five years has several potential benefits that outweigh related limitations. In particular, such nearly back-to-back audits provide necessary information for developing a theory of utilization of program evaluations (see Patton, 1997).

However, the Missouri performance audit report is limited to measuring efficiency for some relatively minor issues in the Division of Tourism. The big issues are not addressed in the 1994 and 1996 audit reports.
For example, the audit report does not address the effectiveness of the Missouri’s tourism advertising program. The 1996 state tourism-advertising budget was $4.7 million, slightly more than 70% of total budget minus the one-time only $1 million library grant. Despite the funds spent on advertising, the report never asks what number of visits to Missouri resulted that would not have been made without these advertising expenditures?

<table>
<thead>
<tr>
<th>Theory in Use:</th>
<th>Metaevaluation</th>
<th>Rationale/additional comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Program evaluation theory used/dedicated by the performance audits?</td>
<td>• Accounting theory of program evaluation</td>
<td>• A highly confrontational audit report and response by the tourism division.</td>
</tr>
<tr>
<td>2. Knowledge displayed in the audit report of availability of alternative theories?</td>
<td>• Other theories of program evaluation not mentioned</td>
<td>• Referring to Solver's, audit report indicates “had” control by the Division</td>
</tr>
<tr>
<td>Classification of Audit:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Comprehensive or limited in scope?</td>
<td>• Limited in scope</td>
<td>• A “...special review”</td>
</tr>
<tr>
<td>2. Focuses on evaluating which of the department activities shown in Fig. 5?</td>
<td>• Focused on one topic</td>
<td>“...we assessed the division’s management controls ...we assessed control risk”</td>
</tr>
<tr>
<td>3. Within the focus of the performance audit, provides thick descriptions of activities done and strategies realized?</td>
<td>• Yes</td>
<td>• E.g., grant made for $1 million by division to corporation before expenses incurred; division’s response: future behavior will be different.</td>
</tr>
<tr>
<td>4. Evaluates research method used by department to measure effectiveness and efficiency of program? Audit conclusion sound?</td>
<td>• Documented analyses; interviewed division personnel</td>
<td>• Most conclusions sound as indicated by agreement and actions of division subsequent to the report.</td>
</tr>
<tr>
<td>5. Evaluates the findings of program impact provided by the department?</td>
<td>• No, Topic not included in this special audit</td>
<td>• Serious problems noted as resolved; some found to still exist. E.g., division overpaid advertising agency $51,000 in 1994; following 1994 audit and finding the Division received $23,000 credit; remainder “attributable to personnel classification errors” according to Division</td>
</tr>
<tr>
<td>6. Reports whether or not this audit is the first, second, or third performance audit done of this department?</td>
<td>• Yes, This is the second audit of the division. First in 1994</td>
<td>• A limited purpose audit</td>
</tr>
<tr>
<td>7. If previous audits done of this department, reviews findings and use of previous audit reports?</td>
<td>• No</td>
<td></td>
</tr>
<tr>
<td>8. Reports scanning done by audit team of tourism-marketing program audits done by other states?</td>
<td>• Partially</td>
<td>• Audit focused on cost controls, efficiency but not effectiveness</td>
</tr>
<tr>
<td>Assessment of Value and Impact of Audit:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Audit covers core issues related to effectiveness/efficiency impacts of activities and results of the program?</td>
<td>• Yes</td>
<td>• $3 million limit in transfer of sales tax revenues violation/recommendation</td>
</tr>
<tr>
<td>2. Audit includes explicitly, concrete, recommendations for change/improvement?</td>
<td>• Highly unfavorable</td>
<td>Auditor’s report: “We will request Attorney General’s opinion regarding the proper interpretation...”</td>
</tr>
<tr>
<td>3. Response by program executives favorable or unfavorable toward recommendations in audit report?</td>
<td>• No</td>
<td>• Issues address program efficiencies only, not effectiveness.</td>
</tr>
<tr>
<td>4. Recommendations in report, if adopted, likely to result in substantial impacts on the quality of activities performed by the department?</td>
<td>• No.</td>
<td></td>
</tr>
<tr>
<td>5. Central issues raised but inadequately addressed in performance audit report?</td>
<td>• No.</td>
<td></td>
</tr>
<tr>
<td>6. Core issues not raised in the audit report?</td>
<td>• Several core issues not raised</td>
<td>• Not raised: scanning; planning; administrating issues; control issues related to effectiveness</td>
</tr>
<tr>
<td>Systems-Thinking and Database Marketing:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Systems thinking displayed in audit report, e.g., policy mapping?</td>
<td>• No.</td>
<td></td>
</tr>
<tr>
<td>2. Report covers implementation of database marketing?</td>
<td>• No.</td>
<td>No evidence of knowledge of systems thinking, policy mapping, database marketing tools.</td>
</tr>
</tbody>
</table>

**Exhibit 7.** Results: Answers for the Issues in a Meta-evaluation of the 1996 Performance Audit of Missouri’s Tourism-Marketing Program and the Division of Economic Development, Division of Tourism.
### Exhibit 8. Key Issues in a Meta-evaluation of the 1996 Missouri Performance Audit. **Note:**

<table>
<thead>
<tr>
<th>Activity/Decision</th>
<th>Did Auditor Collect?</th>
<th>Absence Noted?</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>Accuracy?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scanning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• SWOT assessment?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Market analysis: S/D?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Trends?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Research: customer segmenting?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td><strong>Planning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Written marketing plan?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Evidence of &quot;what if&quot; analyses?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Formal testing of alternatives?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td><strong>Implementing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Milestones established?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Written monitoring guidelines established/followed?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Vital competencies done in-house?</td>
<td>Yes</td>
<td>No</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td><strong>Activity/impact assessing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Realized/planned strategy match?</td>
<td>NA</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Multiple dependent impact measures?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Cost/benefit analysis?</td>
<td>Partial: costs</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• True experiment of marketing impact?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td><strong>Administering</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Quality of written mission statement?</td>
<td>Yes</td>
<td>No absent</td>
<td>Not evaluated</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Annual quantitative objectives established?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Evidence of training, coordination, coaching?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
</tbody>
</table>

**Note:** SWOT = strengths, weaknesses, opportunity, threats; S/D = supply and demand analysis; Levels: L1 = found to be poorly done; L2 = found to be done adequately; L3 = found to be done in superior manner and/or superior performance outcome noted. Accuracy = meta-analysis evaluation of validity of auditor’s conclusion on activity/decision.
The issues that are addressed in the 1996 audit report represent less than $200,000 in expenditures, including the estimated loss of $50,000 in interest on the $800,000 invested of the prepayment of the $1 million library funds. Nearly all of the audit report’s focus relates to potentially unnecessary expenditures and to loss of funds that total less than 3% of the 1996 budget.

The potential public relations disaster of using prisoners and the state Department of Corrections (DOC) for the division’s tourism telemarketing program is not raised in the follow-up section of the 1996 audit report. Would the inquirer knowing that she or he was going to be giving her or his name and address to a prison inmate affect response rates to a tourism-marketing program? Would knowledge of the use of DOC personnel and a legal consultant for such telemarketing work be likely to increase or decrease response rates? How much of an impact would such knowledge have? No information is included in the follow-up section that these issues were addressed.

Exhibits 7 and 8 are summaries of our meta-evaluation of the 1996 Missouri performance audit report on the Division of Tourism. Based on the discussion presented above, our summary evaluation of the Missouri audit report includes the following categories and descriptions:

- coverage of history and budget: excellent
- clarity of writing: excellent
- theory applied: accounting (cost efficiency)
- coverage of issues: limited to minor issues
- style: confrontational
- likely size of impact: little to none

A META-EVALUATION OF A PERFORMANCE AUDIT FOCUSED ON NON-PERFORMANCE OF CORE ACTIVITIES

The 1989 Performance Audit of the North Carolina Department of Commerce, Division of Travel and Tourism

The 1989 North Carolina performance audit is 21 pages including the letter of transmittal written by Edward Renfrow, State Auditor. The following discussion is limited to the portion of the report concerning the Division of Travel and Tourism but not the Film Office.
The letter by Mr. Renfrow implies a comprehensive audit, “The purpose of this audit was to examine the management and operation of the Division of Travel and Tourism and the Film Office and to provide you [the letter is addressed to Governor, Secretary of Commerce, and members of the General Assembly] with our findings and recommendations designed to continue the improvement of those areas.”

The 1989 performance audit report for the North Carolina Division of Travel and Tourism includes a three page executive summary, a half-page of audit objectives, a half-page on audit methods and scope, eight pages of findings, and a three-page letter (dated November 16, 1988) of responses to the audit by Claude E. Pope, Secretary, North Carolina Department of Commerce.

The Executive Summary

The following are the key sentences referring to the Division of Travel and Tourism in the executive summary:

The division was established to promote the growth and development of the state’s travel and tourism industry. The division has an annual budget in excess of $6 million. Over $2 million fund the advertising program that promotes North Carolina as a travel destination. . . . One of the division’s major functions is mailing promotional literature [biggest expense: a visitor’s information guide (VIG)] to people requesting travel information. The division’s conversion study has shown that 51% of those who request information come to North Carolina within the next 12 months and their parties spend an average of $442 while in the State.

Unlike the 1996 Missouri audit report, the 1989 North Carolina audit report gives no detailed information on funding the division or on a schedule of expenses. The conclusion implied by the 51% conversion ratio is that an advertising-and-free-literature combination was the prime influence for the visits and for the subsequent average expenditure of $442. The audit report does not examine details of the conversion study. The audit report does not include questions on the scientific value of the research method used in the conversion study. Telling questions not raised in the North Carolina audit report include:

- what share of inquirers-who-visited North Carolina would have come if they had not been exposed to the advertising and had not inquired?
- what share of inquirers-who-visited NC would have come if they had not received the free literature they requested, or, if they had received a one-sheet, four-sided brochure, rather than a 100+ page VIG?
- what share of non-inquirers who were exposed to the advertising was influenced to visit North Carolina because of the advertising?
- how much additional trip expenditures occurred in North Carolina among visitors exposed to the advertising who received the free literature vs. (a) North Carolina
visitors not known to be exposed to the advertising and who did not request the 
ad-offer of free literature; and (b) North Carolina visitors who were exposed and 
requested, but did not receive, the free literature?

Although the one-group conversion research design (i.e. no control group) used in 
North Carolina conversion studies cannot answer these questions, true-experiment 
designs can, as described by Campbell and Stanley (1963) in program eval-
uation, Banks (1965) in marketing, Caples (1974) in advertising testing, and 
Woodside (1990) in tourism-advertising program evaluation (for details, see 
Exhibit 3).

Because of the main finding on the first page of the executive summary: 
“. . . that the division had failed to answer 93,221 reader service inquiries during 
the first part of 1988,” the final question just raised is of particular interest. (Reader 
service inquiries are requests for ad-offered free literature that are received from 
the reader service postcards, also known as “bingo cards,” in magazines.)

Readers may request information from advertisers by circling the advertiser’s number on the 
postcard. These requests are then forwarded to the division. It has been assumed that reader 
service inquiries are less valuable than the other forms of inquiries (toll-free telephone calls and 
coupons), because less effort is involved. Nonetheless, even if they are assumed to be half as 
valuable, the conversion study suggests that the potential travel revenue associated with these 
unanswered inquiries exceed $10.5 million (North Carolina, 1989, p. 2).

The Department of Commerce responded to this critique by arguing that, “Based 
on the division’s survey of 70 people, the delay in responding had very little, if 
any, impact on their vacation plans” (North Carolina, 1989, p. 3). In reply the 
state auditor argued that:

In our opinion, the critical finding is the division’s failure to take action to remedy the problem 
until it accumulated 93,221 unanswered inquiries. The division’s statement that the delay in 
responding had little short-term effect on vacations is inconsistent with the results of the con-
version study. The conversion study was conducted by the advertising agency retained by the 
department. If that study is inaccurate, that raises the question of whether the state benefits 
from generating and responding to these inquiries (italics added, North Carolina, 1989, p. 3).

The immediate concern is put clearly by the auditor: how can the division both 
claim North Carolina tourist visits due to inquiries generated by the advertising 
and claim no impact due to inquiries generated by the advertising? Unfortunately, 
the even bigger issue is not addressed: no indication is found in the 1989 
North Carolina audit report that the auditor is aware of the more powerful, 
true-experiment designs for estimating advertising-program impact in comparison 
to one-group conversion study designs.

The auditor’s theory that some share of reader service inquiries is associated 
with visits is supported empirically. In a conversion study for Louisiana, 26%
of reader-service card inquirers compared to 36% of direct-response coupon inquirers, and 43% of toll-free call inquirers, converted into visitors (Woodside & Soni, 1988, p. 33).

Two additional points are worth emphasizing here. First, the main impact of the free literature sent to inquirers (i.e. the VIG) is likely to be on affecting the behavioral characteristics of the trip (e.g. on increasing the average expenditures made on the trip, the number of NC regions visited, building word-of-mouth advertising, the overall quality visitor beliefs about the visit, the satisfaction with the total visit experience); the impact of the VIG on generating visits is likely to be positive but much smaller than its impact on the behavioral characteristics of the trip. The combined net profit payoff from creating and providing such detailed information to ad-generated inquirers is likely to be substantial. Unfortunately, while strong evidence is available on the power of VIGs to influence visit characteristics, the theory is not tested via a true-experiment design.

Second, most conversion research studies include sending a sample of inquirers a cover letter and questionnaire that tell the reader that he or she is known to have requested literature. This procedure most likely causes decreases in response rates among inquirers who did not convert into visitors compared to the response rate among inquirers who did convert. Strong empirical evidence supports this proposition (e.g. Woodside & Ronkainen, 1984). The resulting conversion rate, when the sponsor of the study is the state from which the surveyed household member requested information will be very high, often 50% or higher. In comparison, conversion survey studies that do not identify a particular state as sponsoring the study result in lower conversion estimates, for example, 31% overall for the 1987 Louisiana conversion study (Woodside & Soni, 1988).

If the questionnaire and cover letter to refer to travel to several states (e.g. 5–8 southern states), the inquirer receiving the questionnaire does not need to be informed that he or she is known to have requested literature from any one state. Also, additional valuable information on customer intentions and behavior toward competing tourist destinations can be collected in the questionnaire.

The remainder of the executive summary highlights other problems found by the auditor and presents the response by the department. These highlights included the following excerpts.

In the first half of 1988 the division exhausted its supplies of the calendar of events, the accommodations directory, and the state map. As a result 175,000 inquiries received only the attractions study and a form letter stating that the other materials would be sent later. The division failed to provide complete and timely responses to inquiries which potentially could result in some $39 million of travel revenue . . . A major factor leading to the exhaustion of inventories was the division’s failure to reorder publications on a timely basis” (North Carolina, 1989, p. 3).
The following statement is the complete response made by the department, “The division has established a printing schedule” (North Carolina, 1989, p. 3).

The final concern by the auditor was the lack of comprehensive written agreements with industry co-sponsors of business trips to cities outside North Carolina to promote visits to North Carolina. The department’s response to this concern was that, “The division is executing written agreements with the co-sponsors of the travel missions” (p. 4).

Audit Objectives

The performance audit carried out does not reflect the audit objectives set forth in the report. The half-page statement of audit objectives presents the following statement first, “Under the North Carolina General Statues, the State Auditor is responsible for reviewing the economy, efficiency, and effectiveness of state government programs” (p. 5). The largest part of the half-page statement describes the method used to conduct the audit. The audit reports states, “… we reviewed and analyzed: the divisions’ management of its publishing program, including the timeliness of its reorders; the effectiveness and efficiency of the division’s promotional advertising program; and the division’s Welcome Center operation.” The report provides no description of findings or evaluations of the NC welcome center operation.

Audit Methods and Scope

The next section in the report is titled, “AUDIT METHODS AND SCOPE.” The first sentence references the theory frame for the study:

Our performance audits are conducted in accordance with the Standards for Audit of Government Organizations, Programs, Activities, and Functions issued by the Comptroller General of the United States and endorsed by the American Institute of Certified Public Accountants, except that no reviews of general or application controls of the data processing system were completed during this audit (p. 6).

The statement further describes the method used for the audit: review of policies and procedures and interviews with management, staff, and other parties. The last two sentences in this section point out that “This audit does not provide a review of any financial statements of the department. Such a review is incorporated in our comprehensive annual financial report of the State of North Carolina” (p. 6).

The report describes no other theories of program evaluation. The description of the method used for the audit lacks the necessary detail for learning what actions the audit office took in collecting data for the audit. The usefulness of the audit report is restricted severely because the report does not include a review of financial statements of either the department or Division of Travel and Tourism.
The method description indicates that a shallow review was performed to measure the effectiveness and efficiency of the division.

**The Main Body of the Report**
Following the audit methods and scope section, eight pages of description of the activities of the Division of Travel and Tourism provide more details of the problems described above and more responses by the department. These eight pages are entitled, “Chapter 1 – Division of Travel and Tourism.” Page eight mentions that, “The programs described above are in line with the efforts of other states. The vast majority of states use advertising to promote their state to travelers and assist travel writers.” However, no details support this comparison to other states. The chapter includes no benchmarking or specific comparisons to best practices; elsewhere in the report no such comparisons occur.

The main body of the report includes a footnote on page nine to the following statement made in the division’s *Resource Manual* (1988): “The life-blood of the travel industry is response to advertising, and a phenomenal 51% of those who directly answer North Carolina’s ads (i.e. they inquire at this office) travel to the state within 12 months. They (average party of 2.7 persons) spend more than $442 while here.” The footnote:

> These statistics come from the 1984 conversion study done for the division by the advertising agency. The agency surveyed a sample of people who had asked the division for travel information. The survey also provided a demographic profile of the people who are traveling in North Carolina (North Carolina, 1989, p. 9).

The audit report does not raise the issue of lack of independent participation in evaluating the work of the advertising agency. Might the advertising agency be biased in designing an advertising effectiveness study to examine its own performance? Earlier in the audit report the following statement appears: “The division oversees the overall direction of the advertising program. Actual design and implementation of the program is the responsibility of the advertising agency of McKinney and Silver of Raleigh, North Carolina.”

The main body of the audit report identifies lower productivity in the inquiry fulfillment section as the primary cause for not responding in a timely manner to 93,221 inquirers. “A primary cause of the problem has been a fall-off in the productivity of the inquiry section during the first five months of 1988” (North Carolina, 1989, p. 10).

**The Response by the Secretary**
A letter response from Claude E. Pope, Secretary, Department of Commerce, is appended to the audit report. In his letter Mr. Pope mentions that “approximately
70,000” ad-generated inquiries were from reader service cards, and he reiterates that 70 persons were contacted from these cards “to find out what impact the delay in providing the requested material had on their vacation plans for the summer season just ended.” He reports “…very little, if any, effect…” Mr. Pope’s letter includes no details on the follow-up study, such as the following issues:

- How many attempts to interview persons were necessary to result in 70 interviews?
- What problems occur with generalizing results for a population greater than 70,000 based on responses from 70 persons?
- What biases might occur when persons are asked if the information requested and not sent/received affected their summer travel plans?
- What factors account for the difference of more than 23,000 inquiries in the 70,000 reader service inquirers not receiving the requested literature and the 93,221 mentioned in the audit report?
- Why does Mr. Pope make no response on the incomplete response mailing sent to 175,000 additional inquirers?

Mr. Pope makes no reference in his letter to the estimated loss of revenues due to the lack of fulfilling the literature requests.

Mr. Pope writes, “The productivity of the individuals working at the Inquiry Section during the first five months of 1988 was not down; there were two fewer people working at Inquiry during this period than in the previous years” (p. 11). The report does not discuss the possibility of hiring temporary workers to work in the inquiry section. The auditor felt compelled to reply to Mr. Pope’s response: “In our opinion, this response of the department requires a reply. Our concern with the response is its tendency to divert attention from the most critical issue: which is the division’s failure to take action to remedy the problem until the 93,221 inquiries accumulated” (p. 12).


The 1989 North Carolina performance audit identifies substantial numbers of inquiries unanswered and partially answered by the division in 1988. The report includes the suggestion that the estimated loss of these problems exceeds $49 million in travel revenues ($10.5 million + $39 million). Although how much travel revenue was lost is debatable, most likely some amount was lost due to poor coordination of literature fulfillment.

The audit report does not provide estimates of the total value of the state’s tourism-marketing program. Is North Carolina’s tourism-marketing program
profitable (i.e. the taxes net of expenses of the program are a substantial return on the funds invested)? The audit report does not include any specific benchmarking analysis or best practice analysis of government tourism-marketing programs. The report indicates no awareness of theories of program evaluation other than accounting audits, for example, Campbell’s (1969) theory applied to the need for scientific tests of advertising impact and how to conduct such tests are not covered in the report.

The North Carolina audit report presents nothing on the scanning activities or planning activities of the Division of Travel and Tourism. Consequently, a discussion of discrepancies between planning and reaching milestones is limited to the failure to respond to inquiry requests from the advertising program. The audit report describes performing a review and analysis of the division’s welcome center operation. Neither review nor analysis of this operation is included in the report. 

Exhibits 9 and 10 provide a summary of several points in this meta-evaluation of the NC audit report. Based on the discussion presented above, our summary evaluation of the North Carolina audit report includes the following categories and descriptions:

- coverage of history and budget: poor
- clarity of writing: excellent
- theory applied: accounting only
- coverage of issues: limited to a few issues
- style: confrontational
- likely size of impact: little to none
- overall performance quality of report: low

The style of the 1989 North Carolina audit report is confrontational. The report focuses on estimated losses of travel revenues caused by poor coordination in responding to ad-generated inquiries. The response by Mr. Pope, Secretary of the Department of Commerce, conflicts on substantive issues of the size of the problem; he does not respond to the estimated lost revenue impacts. He does attempt to divert attention from the auditor’s findings and minimize problems.

The audit report does not describe activities the division is doing particularly well, if any. The report does not describe recent improvements, if any, in the operations of the division. The report includes no suggestions for improving the division’s scanning or planning activities. The audit report does not discuss the lack of a modern database-marketing program in the division.

The audit report does not discuss the potential public relations and ethical problems of “using women prison inmates to help package responses” (p. 3). Would news stories on this fulfillment program affect the number of inquiries and, therefore, visits to the state?

<table>
<thead>
<tr>
<th><strong>Theory in Use</strong></th>
<th><strong>Meta-evaluation</strong></th>
<th><strong>Rationale/Additional Comments</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Program evaluation theory used/designed by the performance audits?</td>
<td>• Accounting theory of program evaluation not mentioned</td>
<td>• A confrontational audit report and response by the tourism division.</td>
</tr>
<tr>
<td>2. Knowledge displayed in the audit report of availability of alternative theories?</td>
<td>• Other theories of program evaluation not mentioned</td>
<td>• Referring to Scriver, audit indicates &quot;bad&quot; control by the division.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Classification of Audit</strong></th>
<th><strong>Focus is on poor administrating and loss of potential visits due to the failure to fulfill literature requests?</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Comprehensive or limited in scope?</td>
<td>• Limited in scope</td>
</tr>
<tr>
<td>2. Focuses on evaluating which of the department activities shown in Fig. 57 of topics done and strategies realized?</td>
<td>• Focused on a list number</td>
</tr>
<tr>
<td>3. Within the focus of the performance audit, provides thick descriptions of activities done and strategies realized?</td>
<td>• No</td>
</tr>
<tr>
<td>4. Evaluates research method used by department to measure effectiveness and efficiency of program? Audit conclusion sound?</td>
<td>• Partially</td>
</tr>
<tr>
<td>5. Evaluates the findings of program impact provided by the department?</td>
<td>• Audit identifies conflicting assumptions made by the division.</td>
</tr>
<tr>
<td>6. Reports whether or not this audit is the first, second, or third performance audit done of this department?</td>
<td>• Only for inquiries not answered</td>
</tr>
<tr>
<td>7. If previous audits done of this department, reviews findings and use of previous audit reports?</td>
<td>• NA</td>
</tr>
<tr>
<td>8. Reports scanning done by audit team of tourism-marketing program audits done by other states?</td>
<td>• No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Assessment of Value and Impact of Audit</strong></th>
<th><strong>Claim is made that such scanning done but unsupported in the report</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Audit covers core issues related to effectiveness/efficiency impacts of activities and results of the program?</td>
<td>• Partially, focused on effectiveness in program causing revenues</td>
</tr>
<tr>
<td>2. Audit includes explicitly, concrete, recommendations for change/improvement?</td>
<td>• Yes</td>
</tr>
<tr>
<td>3. Response by program executives favorable or unfavorable toward recommendations in audit report?</td>
<td>• Unfavorable</td>
</tr>
<tr>
<td>4. Recommendations in report, if adopted, likely to result in substantial impacts on the quality of activities performed by the department?</td>
<td>• Yes</td>
</tr>
<tr>
<td>5. Central issues raised but inadequately studied; addressed in performance audit report?</td>
<td>• Yes</td>
</tr>
<tr>
<td>6. Core issues not raised in the audit report?</td>
<td>• Limited value of conversion to tourism-marketing cost/benefits of tourism-marketing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Systems-Thinking and Database Marketing</strong></th>
<th><strong>Very little e.g., policy mapping?</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Systems thinking display in audit report, e.g., policy mapping?</td>
<td>• No</td>
</tr>
<tr>
<td>2. Report covers implementation of database marketing?</td>
<td></td>
</tr>
</tbody>
</table>

Note: SWOT = strengths, weaknesses, opportunity, threats; S/D = supply and demand analysis; Levels: L1 = found to be poorly done; L2 = found to be done adequately; L3 = found to be done in superior manner and/or superior performance outcome noted. Accuracy = meta-analysis evaluation of validity of auditor’s conclusion on activity/decision.

<table>
<thead>
<tr>
<th>Activity/Decision</th>
<th>Did Auditor Collect?</th>
<th>Absence Noted?</th>
<th>L1 L2 L3</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scanning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• SWOT assessment?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>• Market analysis: S/D?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>• Trends?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>• Research: customer segmenting?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Planning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Written marketing plan?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>• Evidence of &quot;what if&quot; analyses?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>• Formal testing of alternatives?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Implementing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Milestones established?</td>
<td>Partially</td>
<td>Yes</td>
<td>L1</td>
<td>Accurate</td>
</tr>
<tr>
<td>• Written monitoring guidelines established/followed?</td>
<td>Partially</td>
<td>Yes</td>
<td>L1</td>
<td>Accurate</td>
</tr>
<tr>
<td>• Vital competencies done in-house?</td>
<td>Yes, indirectly</td>
<td>No</td>
<td>L1</td>
<td>Accurate</td>
</tr>
<tr>
<td><strong>Activity/Impact Assessing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Realized/planned strategy match?</td>
<td>Yes</td>
<td>No</td>
<td>L1</td>
<td>NA</td>
</tr>
<tr>
<td>• Multiple dependent impact measures?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td>Accurate</td>
</tr>
<tr>
<td>• Cost/benefit analysis?</td>
<td>Partial: costs</td>
<td>Yes</td>
<td>L1</td>
<td>Accurate</td>
</tr>
<tr>
<td><strong>Administering</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Quality of written mission statement?</td>
<td>Yes</td>
<td>Not evaluated</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>• Annual quantitative objectives established?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>• Evidence of training, coordination, coaching?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td>Accurate</td>
</tr>
</tbody>
</table>

**A META-EVALUATION OF AN AUDIT FOCUSING ON RESEARCH METHOD**

**The 1985 Performance Audit of the Minnesota Tourism Development Program**

development and tourism promotion programs and is funded primarily through the state’s General Fund” (Minnesota, 1985, p. iii).

The Minnesota audit report includes a 12-page executive summary followed by a 236-page report. The second half of the report focuses on the “Iron Range Resources and Rehabilitation Board [IRRRB].” We limit the following analysis to the first half of the report that is focused on DEED. However, several activities of the IRRRB relate directly to tourism-marketing and are described negatively in the audit report; for example, “The IRRRB should halt plans to construct the science museum, the historic village and railroad, and the golf course, and other amenities for the proposed hotel development” (p. xvi).

The following are the main tourism-marketing related findings in the audit report sections focused on DEED:

The department’s expanded tourism promotion efforts appear to have been generally successful in stimulating the state’s economy, although not as successful as program advocates suggested. Our best estimate is that the additional state expenditures on promotion and advertising are generating on average, at least $5 in additional tourism spending per tax dollar spent.

The effectiveness of the 1985 television campaign needs to be closely monitored to ensure that future promotional expenditures are targeted to those markets and media that bring the highest return. We also recommend that the Minnesota Office of Tourism develop a methodology for estimating the return on investment for major advertising campaigns, particularly those conducted outside Minnesota. The office has in place some of the studies and surveys needed to estimate the return that the state is receiving for its tax dollars. However, the office needs to combine survey results with a methodology similar to but more extensive than the one we used in this report. We believe it is important to know which of the various expanded tourism promotional efforts have been worth their costs. (p. xii)

Note that this statement in the Minnesota audit report is a more sophisticated view (i.e. it uses a theory of tourism-marketing performance auditing) than found in the Missouri (1996) and North Carolina (1989) reports. Focusing on estimating the return on investment of tax dollars of tourism-marketing expenditures sets the Minnesota report apart from the Missouri and North Carolina report. However, note that the quotation above does not estimate net tax dollars generated by the promotion and advertising expenditures but rather estimates $5 in additional tourism spending per tax dollar spent.

Chapter 1 in the report starts with a background section. This section describes the state’s approval of a $6.1 million increase in the Office of Tourism’s biennial budget for use in attracting additional tourists to Minnesota. This increase was one outcome of “a severe economic recession and increased competition among states for new business development . . .” (Minnesota, 1985, p. 1).

The second section of Chapter 1 in the report implies that the audit is comprehensive, not limited in scope. “Instead, this report focuses on whether
economic development programs have been designed and administered so that the state obtains the maximum benefits possible.” The report does not describe stakeholders other than the state.

The audit report includes neither a method section nor does the report refer to a particular grounding in theory of program evaluation. However, the numerous tables on revenues and costs, and the ROI discussions, imply an accounting and financial foundation for the study.

The budget data in Table 2.1 of the audit report indicate substantial increases in tourism-marketing expenditures:

<table>
<thead>
<tr>
<th>Year</th>
<th>Budget in $ Millions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1983</td>
<td>$1.2</td>
</tr>
<tr>
<td>1984</td>
<td>$3.9</td>
</tr>
<tr>
<td>1985</td>
<td>$4.7</td>
</tr>
</tbody>
</table>

The following theoretical assumptions and audit issues conclude Chapter 2 of the report. “It is believed that with the increased awareness about Minnesota as a vacation area, more Minnesotans will vacation in the state and more tourists from other states and countries will be attracted to Minnesota, thus stimulating economic activity for tourism-related businesses” (Minnesota, 1985, p. 15). “In reviewing these efforts [DEED’s tourism-marketing program], we asked:

- How successful are promotional activities in stimulating economic activity?
- How much benefit is the state receiving from its investment in promotion?
- Are the programs targeted to audiences that are likely to be influenced to consider Minnesota in their vacation or business plans?” (p. 15)

Note that the issue of target audiences is a sophisticated question not raised in the Missouri and North Carolina audit reports. Given that performance may vary substantially among different tourism-marketing activities, such a question indicates thoughtful sensemaking by the Minnesota auditors.

Chapter 6, “Tourism Promotion,” in the Minnesota (1985) report is the main section on the state’s tourism-marketing program. Chapter 6 includes “Rationale for State Tourism Promotion;” this section describes two theories: one supporting tourism promotion and one opposed. The bottom-line reason for advocates of tourism-marketing is that “The net increase in spending within Minnesota [caused by tourism-marketing] will stimulate the state’s economy.”
A negative reason for tourism-marketing expenditures is that opponents “question whether benefits gained at the expense of other states will last. Increasing Minnesota’s tourism promotion budget may cause other states competing for the same tourists to increase their budgets. The net effect might not be beneficial to any of the states because the impact of their promotional campaigns would offset one another” (Minnesota, 1985, p. 87). The theoretical causal chain in this argument assumes implicitly that tourism-marketing does not affect the total tourism-revenue pie, that all states have benefits equally attractive for tourists, and that all states have equal abilities in spending tourism-marketing funds. All three of these implicit assumptions are likely to be inaccurate.

The report poses an important concern, “...we believe it is useful to focus the debate on whether increased tourism promotion has produced significant benefits for Minnesota and whether further increases [in tourism-marketing expenditures] are in the best interests of the state” (Minnesota, 1985, p. 88). What is valuable in particular about the Minnesota audit report is that it describes two sets of assumptions (i.e. two theories) on the impacts of tourism-marketing programs. Posing such a debate indicates that the auditors do not assume one view is always accurate – a valuable stance for achieving improved sensemaking.

The following comments are a brief summary of the two auditors’ central observations and findings. First, the auditors examined the tourism office 1982 conclusion that the state tourism promotion returned $43 in benefits for each dollar invested during fiscal year (FY) 1981. After developing a “modified inquiry-conversion method,” and after calculating “The average return to tourism advertising during FY 1984... by dividing tourism expenditures of $32.4 to $36.4 million by costs of $2,270,000, the auditors conclude that the estimated average return is between $14.30 and $16.00.”

The audit report describes “several major limitations” in the previous study that resulted in the 43:1 return on advertising expenditures. These limitations are summarized here; they include the following points (Minnesota, 1985, p. 91):

- the conversion study did not measure the marginal return from the increased state spending on tourism promotion;
- the study failed to consider the effect of non-response bias and, thus, substantially overestimated the percentage of persons who vacationed in Minnesota;
- the results are based on responses to a single mailing;
- the study assumed that all persons who said they had vacationed in Minnesota had come as a result of the office of tourism promotion campaign;
- the study did not consider other factors that would have resulted in a higher estimated rate of return; for example, the study did not attempt to measure how many people vacationed in Minnesota as a result of the state’s
advertising campaign even though they did not request information from the tourism office.

The audit team applied conversion survey data for 1984 to estimate tourism expenditures. "After each advertising season, the tourism office surveys a sample of persons who telephone or write to the office and request tourism brochures" (Minnesota, 1985, p. 111). The audit team did its own data analysis of the surveys. "Tourism expenditures were calculated by multiplying the following factors together:

- the number of inquires;
- the percentage of those surveyed who vacationed in Minnesota;
- the percentage of those surveyed who said that the brochure helped them decide to vacation in Minnesota;
- the average amount those surveyed said they spent during their vacation in Minnesota" (Minnesota, 1985).

Although this method has improvements over almost all tourism-marketing conversion studies; nevertheless, it includes telling weaknesses compared with the use of a true experiment design with advertising treatment and control groups. Among others, these weaknesses include the following points:

- assuming that respondents are able to know whether or not the brochure helped them decide (a leading question);
- the biases in asking only about travel to Minnesota and reminding the survey respondents that they are known to have requested the brochure;
- the inability to measure the impact of the advertising on persons exposed to it who did not request the brochure;
- and, not adjusting for the over-estimate in conversion caused by the non-response bias, due in part to the use of a single mailing of the questionnaire.

The audit includes a section on “planning and research activities” of the office of tourism. The report includes giving credit to the office for engaging in scanning and planning activities (pp. 102–103). However, the audit report does not state whether or not the office of tourism prepares a written marketing plan. The report does point out improvements made by the office in using conversion research to measure advertising effectiveness: “Beginning in 1983, the tourism office reduced non-response bias by using two [survey] mailings instead of one mailing. Since the [tourism research] literature indicates that non-response bias for surveys with two mailings can still be significant, the tourism office needs to do more analysis of non-response bias to improve accuracy of the survey results” (Minnesota, 1985, p. 104).
Meta-evaluation of the 1985 Minnesota Performance Audit Report

The 1985 Minnesota tourism-marketing performance audit focuses on valuing results from a post positivism Scriven (1980) perspective. The auditors spent the majority of their efforts on improving the estimation of the return on investment of the state’s tourism-marketing program. However, they do devote a little reporting space to describe briefly the scanning, planning, and implementing activities of the state’s Office of Tourism. The audit report is less confrontational than the Missouri (1996) and North Carolina (1989) reports. Surprisingly and disappointingly, the Minnesota audit does not contain a response from DEED.

<table>
<thead>
<tr>
<th>Activity/Decision</th>
<th>Did Auditor Collect?</th>
<th>Absence Noted?</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scanning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• SWOT assessment?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Market analysis: S/D?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Trends?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Research: customer segmenting?</td>
<td>No</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Written marketing plan?</td>
<td>?</td>
<td>No</td>
<td>L2</td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Evidence of “what if” analyses?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Formal testing of alternatives?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Implementing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Milestones established?</td>
<td>?</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Written monitoring guidelines</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>established/followed?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Vital competencies done in-house?</td>
<td>Yes</td>
<td></td>
<td>L2</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>Activity/Impact assessing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Realized/planned strategy match?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>measures?</td>
<td>Yes</td>
<td></td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Multiple dependent impact</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administering</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Cost/benefit analysis?</td>
<td>Partially</td>
<td>Problems L1</td>
<td></td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• True experiment of marketing</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>impact?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Exhibit 12. Key Issues in a Meta-evaluation of the 1989 Minnesota Performance Audit of the Division of Travel and Tourism. Note: SWOT = strengths, weaknesses, opportunity, threats; S/D = supply and demand analysis; Levels: L1 = found to be poorly done; L2 = found to be done adequately; L3 = found to be done in superior manner and/or superior performance outcome noted. Accuracy = meta-analysis evaluation of validity of auditor’s conclusion on activity/decision.
As described earlier, the Minnesota performance audit of the Office of Tourism’s activities and impacts raises more sophisticated issues than found in the Missouri and North Carolina audit reports. However, the auditors do not indicate knowledge of the literature on true experiment designs in marketing and advertising. Consequently, their recommendations for estimating effectiveness and efficiency impacts are limited in value; the recommendations do not illustrate the paradigm shift to the use of true experiments.

Exhibits 11 and 12 provide summaries of our meta-evaluation of the 1985 Minnesota audit report of the state’s tourism-marketing program. Based on the discussion presented above, our summary evaluation of the Minnesota audit report includes the following categories and descriptions:

- coverage of history and budget: useable
- clarity of writing: excellent
- theory applied: accounting only
- coverage of issues: limited: one issue in depth
- style: non-confrontational
- likely size of impact: modest
- overall performance quality of report: good, steps in right direction.

A META-EVALUATION OF A NON-AUDIT AUDIT

The 1995 Tennessee Performance Audit Report of the Department of Tourist Development

This 13-page report includes a transmittal letter (dated July 28, 1995) to several members of the Senate and House of Representatives, State of Tennessee, from W. R. Snodgrass, Comptroller of the Treasury. In his letter, Mr. Snodgrass states, “This report is intended to aid the Joint Government Operations Committee in its review to determine whether the department should be continued, abolished, or restructured” (Tennessee, 1995, p. i).

Purpose and Authority for the Audit

Several sentences early in the report imply that the 1995 Tennessee performance audit was intended to be a comprehensive study. “The objectives of the audit were to review the department’s legislative mandate and the extent to which the department has carried out that mandate efficiently and effectively and to make recommendations that might result in more efficient and effective accomplishment of the department’s legislative mandates” (p. ii). The report gives four objectives for the audit. These objectives include:
to evaluate the efficiency and effectiveness of the department’s administration of its programs;
to develop recommendations, as needed, for department or legislative actions that might result
in more efficient and/or more effective operation of the department.

“Under Section 4-29-217 of that statute, the Department of Tourist Development
is scheduled to terminate June 30, 1996 . . . . This audit is intended to aid the Joint
government Operations Committee in determining whether the Department of
Tourist Development should be abolished, continued, or restructured” (Tennessee,

The 1995 Tennessee audit includes the following sections:

* Purpose and Authority for the Audit (1/4 page);
* Objectives of the Audit (1/4 page);
* Scope and Methodology of the Audit (1/4 page);
* Organization and Statutory Duties (1(2/3) pages);
* Economic Impact of Tourism (1/3 page);
* Observations and Comments (1(2/3) pages);
* Findings and Recommendations (4 pages), includes response comments.

In the scope and method section, the report states that, “The current activities
and procedures of the department were reviewed, focusing on procedures and
conditions in effect during field work, May through July 1994. The audit was con-
ducted in accordance with generally accepted auditing standards and included . . . .”
The method reported included: (1) review of statues, regulations, policies, and
procedures; (2) document analysis of departmental files; (3) interviews with a
“contracted vendor,” directors of regional tourist associations, and staff members;
(4) site visits to the 12 welcome centers (p. 1).

Comments on the Method Described in the Report
The report includes no details of the method used. No findings on the policies and
procedures are described. No details on how many documents were reviewed;
the contents of documents reviewed are not discussed in the report. Who was
interviewed and how many persons were interviewed are not reported. The details
on the data collected during the welcome center visits are absent.

A Review of the “Organization and Statutory Duties” Section of the Report
The report states, “For the year ended June 30, 1994, the department had expendi-
tures of $9.3 million and revenues and appropriations of $9.5 million . . . . During
fiscal year 1994, the department had 152 staff positions, 108 of which were at
the welcome centers” (p. 2). Neither this section nor any other section of the
report includes financial statements. The number of staff positions appears
excessive but this issue is not raised; the welcome center operation includes
five regional managers: one each for the northeast, middle, west, southeast, and middle east.

The Main Information in the Report
The heart of the report appears on 1/3 of a page; the section is “Economic Impact of Tourism.” This section makes the following points without supporting evidence presented anywhere in the report:

The department collects and analyzes data on the economic impact of tourism on the state. “According to the department, the 38.9 million tourists traveling to and through Tennessee in 1993 spent $6.77 billion that generated $517 million in state and local taxes. . . . In 1992, approximately 460,000 people visited Tennessee from Canada and nearly 2,000,000 visited from overseas countries” (Tennessee, 1995, p. 5).

Main Findings in the Report
The report describes two findings. First, “The cost-effectiveness of using state personnel instead of outside contractors should be analyzed. Since January 1994, the department has contracted with Wessan Interactive Network of Omaha, Nebraska, to answer the state’s 800 tourist information number, mail requested information, and provide a summary of caller characteristics” (p. ii). “The Department of Tourist Development conducted no formal study and gathered no empirical data to support the need to go outside state government for these services” (p. 7).

Second, “the Department needs to ensure that staff comply with water quality statues and rules” (p. 11). This finding relates to hygiene management of the 12 welcome centers.

We offer the following observations about the first finding. An organization cannot achieve the objective of getting close to its customers by contracting out contacts with customers. Effective marketing practices requires direct interactions with customers. Nuances about customer beliefs, preferences, and behaviors are learned by talking with them. This principle about getting and staying close to customers is described in the strategic management literature (e.g. see Peters, 1987; Peters & Waterman, 1982). Both Tennessee auditors and the state’s tourism executives appear to be unaware of the principle. Contracting out answering customers’ telephone calls is likely to be more efficient initially, but it is also less effective than building an in-house direct-customer contract program.

To use an analogy often found in the Economist, the second finding is a “small beer” compared to major issues that are not addressed in the report. For example, the findings in the report do not include a discussion of the need for database marketing. The findings do not discuss effectiveness issues related to developing department relationships with inquirers. Given that the high value in establishing
and building relationships with prospects and customers is a central principle of effective marketing strategy (e.g. see Kotler, 1997), then several more powerful reasons can be made for answering inquiry telephone calls, sending literature, and communicating with customers from Nashville, not Omaha.

Potential “cost-effectiveness” (really, cost-efficiency) is one reason for bringing the literature-fulfillment program home to Tennessee, especially given the initial likely higher efficiency expertise of Wessan Interactive Network. However, for reasons related to effectiveness (e.g. revenues, visit quality, customer satisfaction with service encounters), Tennessee should perform the operation by full-time Department staff members.

The “Management Comment” to the first main finding was that, “The department believes there is adequate information available to justify its decision to establish the Wessan contract. All state procedures were followed in the award of the contract” (p. 9). Surprisingly, the audit report includes no attempt to actually evaluate the “cost-effectiveness of contracting these services” but simply states such an evaluation should occur.

**Meta-evaluation of the 1995 Tennessee Performance Audit Report**

Exhibits 13 and 14 summarize the details of our meta-evaluation of the 1995 Tennessee performance audit. The 1995 Tennessee performance audit is unique in its lack of measurement and evaluation of the activities and impacts of the state’s tourism-marketing programs. Almost no information is provided on the scanning, planning, implementing, and impact assessment, and administering activities of the Department of Tourist Development. The audit report expresses a possible problem with using an outside contractor for the state’s literature-fulfillment program because the department did not evaluate the possibility this work could have been done in Tennessee with lower costs. The report does not describe why the auditor did not conduct such a cost-analysis evaluation.

However, we do not mean to suggest that this cost-efficiency issue is a valuable one to be resolved. As described above, other, more powerful, reasons are available for answering customer inquiries in Tennessee, from Nashville, and in the offices of the department. Also, we estimate that the potential cost savings of running the existing fulfillment program represent less than one percent of the department’s annual budget. Also, the cost savings of Tennessee vs. Wessan fulfillment programs are likely to be negative, because of the substantial experience-curve impact (i.e. cost-then-price reductions based on accumulated fulfillment experience) in a professional fulfillment organization, such as Wessan Interactive Network.
### Theory in Use:

<table>
<thead>
<tr>
<th>Meta-evaluation</th>
<th>Rationale/Additional comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Program evaluation theory</td>
<td>Accounting theory of program evaluation</td>
</tr>
<tr>
<td>2. Knowledge displayed in the audit report of availability of alternative theories?</td>
<td>A mildly confrontational report and response by the tourism division cited</td>
</tr>
</tbody>
</table>

### Classification of Audit:

| 1. Comprehensive or limited in scope? | Comprehensive in scope |
| 2. Focused on evaluating which of the economic department activities shown in Fig. 5? | Focused on possible cost, impact of advertising |
| 3. Within the focus of the performance audit, provides thick descriptions of activities the done and strategies realized? | Savings in fulfillment |
| 4. Evaluates research method used by department to measure effectiveness and efficiency of program? Audit conclusion sound? | No |
| 5. Evaluates the findings of program impact provided by the department? | No |
| 6. Reports whether or not this audit is the first, second, or third performance audit done of this department? | No |
| 7. If previous audits done of this department, audits reviews findings and use of previous audit reports? | NA |
| 8. Reports scanning done by audit team of tourism made by marketing program audits done by other states? | No |

### Assessment of Value and Impact of Audit:

| 1. Audit covers core issues related to net effectiveness/efficiency impacts of activities and results of the program? | No |
| 2. Audit includes explicit, concerns, not recommendations for change/improvement? | Only minor topic recommendations |
| 3. Response by program executives favorable or unfavorable toward recommendations in audit report? | Unfavorable |
| 4. Recommendations in report, if adopted, likely to result in substantial impacts on the quality of activities performed by the department? | No |
| 5. Central issues raised but inadequately addressed in performance audit report? | No |
| 6. Core issues not raised in the audit report? | Yes |

### Systems-Thinking and Database Marketing:

| 1. Systems thinking display in audit report? | None |
| 2. Report covers implementation of database marketing? | No |

---

**Exhibit 13.** Results: Answers for the Issues in a Meta-evaluation of the 1989 Performance Audit of Tennessee’s Tourism-Marketing Program.
Assessing Alternative Methods of Performance Evaluation

<table>
<thead>
<tr>
<th>Activity/Decision</th>
<th>Did Auditor Collect?</th>
<th>Absence Noted?</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>Accuracy?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scanning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• SWOT assessment?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Market analysis: S/D?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Trends?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Research: customer segmenting?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td><strong>Planning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Written marketing plan?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Evidence of ‘what if’ analyses?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Formal testing of alternatives?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td><strong>Implementing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Milestones established?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Written monitoring guidelines established/followed?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Vital competencies done in-house?</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td><strong>Activity/Impact assessing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Realized/planned strategy match?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Multiple dependent impact measures?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Cost/benefit analysis?</td>
<td>No</td>
<td>Yes (Possible problem noted)</td>
<td></td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• True experiment of marketing impact?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td><strong>Administering</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Quality of written mission statement?</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Annual quantitative objectives established?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Evidence of training, coordination, coaching?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
</tbody>
</table>

Exhibit 14. Key Issues in a Meta-evaluation of the 1989 Tennessee Performance Audit of the Division of Travel and Tourism. Note: SWOT = strengths, weaknesses, opportunity, threats; S/D = supply and demand analysis; Levels: L1 = found to be poorly done; L2 = found to be done adequately; L3 = found to be done in superior manner and/or superior performance outcome noted. Accuracy = meta-analysis evaluation of validity of auditor’s conclusion on activity/decision.

The Tennessee performance audit may be prototypical of a non-audit audit. The performance of such a minimal audit may be particularly disappointing given the circumstances stated early in the report. Stated again: “This report is intended to aid the Joint Government Operations Committee in its review to determine whether the department should be continued, abolished, or restructured” (p. i). The report offers no evidence that the activities of the department lead to tourism activities, expenditures, or net tax profits for the state that would not have occurred if the department had not existed. Consequently, the report does not provide evidence in support of continuing the department; the report does not address this issue directly in its findings – another indicator of the report’s uniqueness as a non-audit audit.
Based on the discussion presented above, our summary evaluation of the 1995 Tennessee audit report includes the following categories and descriptions:

- coverage of history and budget: poor
- clarity of writing: good
- theory applied: accounting (cost efficiency)
- coverage of issues: limited to minor issues
- style: mildly confrontational
- likely size of impact: little to none
- overall performance quality of report: very low.

A META-EVALUATION OF A NEGATIVE COMPREHENSIVE AUDIT

The 1987 Hawaii Management Audit Report of the Hawaii Visitors Bureau and the State’s Tourism Program

This 1987 Hawaii performance audit report is comprehensive in covering the topics of scanning, planning, implementing, measuring results, and administering by the Hawaii Department of Planning and Economic Development (DPED) and the Hawaii Visitors Bureau (HVB). Following Scriven’s (1980) theory of evaluation, the 1987 Hawaii audit is a prototype of a very good audit indicating very bad performance overall by both the DPED and the HVB. The conclusion that the quality of the 1987 Hawaii audit report is substantially better than the 1995 Tennessee audit understates the differences in quality of the two reports.

Primarily because of historical reasons, the HVB performs most of the decisions and actions for Hawaii’s tourism-marketing program, a private non-profit organization. Funding to run the HVB comes mostly from the state of Hawaii, “Today, legislative appropriations account for nearly 80% of HVB’s operating budget” (Hawaii, 1987, p. 33). Legislative statues mandate that tourism-marketing funds be appropriated to the state agency charged with tourism development, DPED, “that agency would then enter into a contract with HVB for the promotion and development of tourism” (p. 33).

During the 1970s and 1980s the Hawaiian Legislature passed a number of statues to get the DPED to take charge of directing the state’s tourism-marketing program and requiring that the HVB follow the directives of the DPED. Substantial evidence is found throughout the 1987 audit report that both the DPED and HVB have ignored these statues. The report concludes that much effort and
time are spent attempting to communicate and coordinate actions between DPED and HVB administrators and staff, the resulting quality of managing Hawaii’s tourism-marketing program is very low.

The bureau [the HVB] enjoys the operating freedom of a private organization with virtually guaranteed, substantial state funding but with no need to produce profits, unhampered by the reviews and controls of a regular state agency, and accountable to no one for organizational effectiveness. As pointed out in Chapter 4, the Department of Planning and Economic Development (DPED) has not been aggressive in enforcing its contract with HVB. Theoretically, the State could contract with any other organization or organizations for all or part of its tourism marketing services. At one time, the Legislature directed DPED to contract directly for tourism advertising. But over time, HVB has convinced state decision makers that it is the best entity to market tourism for the State. Except for the review of its annual budget request at the Legislature, HVB is not held to account (Hawaii, 1987, p. 97).

The Contents of the 1987 Hawaii Performance Audit Report

This audit report contains 221 pages. The audit is organized into ten chapters followed by comments (letter plus 4 pages) by Roger A. Ulveling, Director of the DPED, and by comments (17 pages in a letter) by Walter A. Dods, Jr. (HVB Chairman of the Board) and Stanley W. Hong (HVB President).

Chapter 1: Introduction (2 pages). Chapter 1 states three objectives for the audit: examine and evaluate the role of the DPED in managing the State’s tourism program; assess the effectiveness of the programs and operations of the HVB; and to make recommendations.

Strong evidence is found in the audit report to support the auditor’s view that, “The audit took a broad approach to the industry and its place in Hawaii’s economy. We reviewed the trends in worldwide tourism because Hawaii competes in a global economy.” This claim is supported with a thorough literature review and references in later chapters in the report.

Chapter 2: The Tourist Industry Worldwide (12 pages). This chapter includes benchmarks and best practices data in government tourism-marketing programs. For example, the chapter describes the New Zealand’s database-marketing program. “The New Zealand tourist office keeps a computerized list of the 50,000 to 100,000 inquiries it receives each year as a result of its mail-in advertising coupons. The list is then used for direct mail promotion and is also provided to travel operators and agents” (Hawaii, 1987, p. 10, quoted from Nick Verrasto in ASTA Travel News, June 15, 1986, p. 25). While not noted in the audit report, Hawaii does not have a similar database tourism-marketing program – one indication of the lack of systems-thinking that is widespread in government tourism-marketing programs in all U.S. states.
Chapter 2 includes a report on trends in travel products and markets. For example, “To stimulate demand, many areas are creating new travel products and purposes. These cater to special interests or hobbies, e.g. special hiking tours or adventure trips” (Hawaii, 1987, p. 14). Chapter 2 includes a total of 34 footnotes to data and literature on tourism-markets.

A description of the Canadian-government study of U.S. travelers supports the proposition that “Market researchers recognize that travelers are not a homogeneous group” (Hawaii, 1987, p. 9). While seemingly obvious, this proposition is important to support empirically. The proposition leads to several related, less obvious propositions, such as:

- for a given destination, different visitor segments seek different destination experiences;
- some prospective visitor segments are not worth spending government tourism-marketing funds to attract because these segments do not want to visit or if they did, they will not spend enough money to justify attracting.

Chapter 2 earns only a “good,” not a “very good,” rating because the chapter does not describe nor identify specific competitors with Hawaii. Consequently, the unique features of Hawaii products-markets are not compared to the competitor’s unique products-markets. Still, recognizing the need for such coverage as the contents of Chapter 2 goes well beyond that found in any state’s performance audit of tourism-marketing programs.

Chapter 3: Tourism in Hawaii (19 pages). Chapter 3 reports a history of tourism behavior in Hawaii, including benefits and costs. “In 1986, there was a 14.8% increase in the number of visitors over 1985 with total expenditures estimated at $5.8 billion. Given the multiplier effect of tourist expenditures, the visitor industry ultimately contributes to over half of the gross state product and supports more than one in three jobs in Hawaii” (p. 19).

The chapter describes some costs. On the negative side, the auditors perceive that substantial profits from tourism leave the state and benefit only the multinational corporations that control more and more of Hawaii’s tourist industry. “One economist estimates that almost half of the $4 billion that was spent in Hawaii in 1984 left the State almost immediately” (p. 21). Two points are clear from a 1986 survey of Hawaiian residents: (1) they agreed that the positive economic benefits of the [tourism] industry outweighed the negative social impacts; and (2) an overwhelming majority felt that the economic benefits were not more important than protecting the environment.

Chapter 3 reports detailed profiles of Hawaii’s westbound visitors (i.e. from California and other states) and eastbound visitors (i.e. mainly from Japan). Here
is one comparison: Japanese tourist parties are estimated to spend $250 per day in Hawaii compared to U.S. tourist parties who are estimated to spend $98 per day (Hawaii, 1987, p. 23).

Chapter 3 describes government funding of tourism-marketing and related programs. The point is made that for the first 40 years of its existence (since 1903) the HVB received $1 in government support for every $2 it received in [membership] subscriptions from the business community. However, the ratio changed after World War II. “Today, legislative appropriations account for nearly 80% of HVB’s operating budget” (p. 33).

During the 1970s the state developed a state plan for tourism. “The Hawaii State Plan, codified as Section 226, Hawaii Revised Statues, was enacted in 1978. It establishes the overall theme, goals, objectives, and policies for the State and a state planning process” (p. 34).

Wow! States rarely have a plan for tourism let alone a tourism plan enacted into law. “The State Tourism Functional Plan was approved by the Legislature in 1984 . . . State government is responsible for administering the objectives and policies in the functional plan” (p. 35). A “tourism branch” in the DPED is currently responsible for coordinating activities in implementing the state’s tourism plan. The head of the tourism division reports to one of the two deputy directors at DPED.

The branch has three full-time positions: a tourism-program officer, a tourism specialist, and a secretary. The branch sees itself as having two main responsibilities. The first is the administration of the State’s contract with HVB including preparing and assisting in negotiating the annual contract. It sees its responsibility in this area as primarily fiduciary. Branch personnel report that they attend many of the HVB committee meetings to keep on top of HVB activities (p. 35).

What is implied here becomes clear after reading the rest of the report: state government in Hawaii has organizationally buried the governance and leadership of its tourism plan and the state’s tourism-marketing responsibilities. Also, the allocation of two government professional positions for tourism managing responsibilities is a further indicator that, to a great extent, the state has abdicated its participation in statewide tourism-marketing programs.

Consequently, in the following chapter, the auditor recommends that “An office of tourism be established in the Department of Planning and Economic Development that is headed by a deputy director . . . The office should also strengthen the State’s role in budgeting, including the restoration of the tourism program as a separate and identifiable program in the executive budget” (Hawaii, 1987, p. 69). Given the substantial number of statutes passed by the Hawaiian government requiring the state itself to manage its tourism programs, including the state’s tourism-marketing programs, and given the serious shortcomings in the marketing efforts of the HVB reported by several sources, the Director of
the DPED might be expected to support the auditor’s recommendations for an Office of Tourism and a Deputy Director of Tourism. However, Mr. Ulveling and the DPED disagreed: “It [the response by Mr. Ulveling and the DPED] does not agree with our recommendation that an office of tourism be established. It also does not see the need to maintain the State Functional Plan for Tourism” (Hawaii, p. 197).

Close reading of DPED’s response to the management audit strongly supports the conclusion that the state has abdicated its leadership and management responsibilities in tourism. For example, the DPED response includes the following statement in response to the recommendation that an Office of Tourism and Deputy Director position be created in the DPED:

The report comments unfavorably on the current DPED economic development organization. The implication is that creation of a branch within a line division rather than the former staff office is detrimental to the tourism mission of the department. While it may appear that the change in reporting level decreases visibility and departmental support, in fact the opposite is true. With the addition of a division head and staff, the tourism branch now has more people available to assist and support it than ever. The division head and professional staff backs up the branch manager in preparing various documents, attending meetings, and administering contracts (Hawaii, 1987, p. 203).

This response does not support the view that DPED is responsible for leading and managing the state’s tourism-marketing programs. The response does not suggest that tourism behavior and industries are central to the economic well being of the state. Having the state’s senior, full-time, head of tourism operations be branch manager who reports to a division head, who reports to a deputy director, who reports to the director of DPED, more than suggests that the DPED has decided to be minimally involved in tourism. Note that the branch manager’s role is described to be one of “preparing various documents, attending meetings, and administering contracts.” The DPED’s response indicates little responsibility for senior management activities regarding tourism-marketing. Also, note that in 1987 the head of this three-person branch (including a secretary) administers the contract with the 89-member staff of the HVB.

Chapter 4: The state’s tourism program (32 pages). Poor performance of the DPED in all areas of responsibility is the main finding in Chapter 4.

We find the following: (1) Despite repeated studies and agreements over the years about the State’s goals, objectives, and policies for tourism, there has been little consideration or concerted efforts towards attaining these objectives. (2) Although the Department of Planning and Economic Development (DPED) is the State’s lead agency for tourism, the State has yet to implement a tourism program because the department has played a reactive and passive role. There is an absence of leadership and focus to its efforts. (3) The department has yet to clarify its responsibilities for the tourism program vis-a-vis those of the Hawaii Visitors Bureau.
Chapter 4 clearly indicates agreement “that the State has a broader role to play than HVB. The State is responsible for the entire tourism program. This [responsibility] would include coordinating infrastructure needs, monitoring the industry and its impact on the community [and the environment], and ensuring the [high] quality of the visitor industry generally” (Hawaii, 1987, p. 39).

Several tourism studies conducted, and then ignored by the State of Hawaii, are described briefly in Chapter 4. One study in particular is eye catching, a 1970 study commissioned by DPED with Mathematica of Princeton, New Jersey, to assess the costs and benefits of Hawaii’s tourist industry. Mathematica is a firm noted for its expertise in policy mapping and systems dynamics modeling. Its work for Hawaii is one indicator supporting the proposition in Chapter 4 that “The State has conducted numerous studies of considerable value. For the most part, however, the studies have been ignored. They have not been used productively to build on the work that has gone on before” (Hawaii, 1987, p. 39).

Chapter 4 reviews several reports by the DPED and details of the “State Tourism Functional Plan” passed by the Legislature supports the view that the public sector is the lead organization to administer State objectives and policies of the State Tourism Functional Plan. However:

Despite this long history of repeated demands for a larger government role and the existence today of an agency that says it is responsible for being a lead organization for tourism, the same complaints are still being heard . . . . The State appears to have gone full circle. The same complaints that were being made in 1957 are still being heard today. The need for coordination remains. Infrastructure continues to be a problem. The impact of the visitor industry on the State is of continuing concern. And government is still being prodded to assume its responsibilities for tourism (pp. 47–49).

Specific lack of managing activities (see Fig. 5) by the DPED are reviewed in Chapter 4. For example, “DPED has not budgeted funds for tourism promotion or for a tourism program. Instead, DPED merely forwards HVB’s budget requests without analysis [to the Legislature]. The department testifies before the Legislature on HVB’s budget request, but there is no analysis of whether the amounts requested for promotion are insufficient, adequate, or too much” (p. 51).

The following false view (i.e. theory of advertising program evaluation) expressed by the 1986 Director of the DPED, Kent M. Keith, to the Legislature reflects the primary cause for the lack of high quality studies on advertising effectiveness of tourism-marking programs. Note that in its response to Mr. Keith’s propositions, the audit report identifies the solution, even though the audit
report does not describe the availability of scientific tests to measure the impact of advertising and other marketing actions.

In 1985, the director [of the DPED, Kent M. Keith] commented in his testimony that while it was impossible to estimate the share of the additional $300 million increase in visitor expenditures for 1983 for which HVB was responsible, even a relatively small share would have handsomely repaid the State’s investment. The director believed that the investment was repaid many times over.

This kind of testimony is of little assistance to legislative decision makers. It reflects DPED’s general lack of concern [about] or information on the impact of funds appropriated to HVB and the effectiveness with which they are expended. The department still has not developed a system for oversight of HVB nor has it identified any more useful measures of effectiveness. Consequently, it has no means for assessing existing promotional efforts or evaluating potential new markets (Hawaii, 1987, pp. 51–52).

The rest of Chapter 4 provides evidence supporting the conclusion that DPED is, at best, incompetent. The topics covered include the following items. “Failure to take actions relating to physical [tourism infrastructure] development” (p. 53), “Failure to take actions relating to employment and career development” (p. 54), “Failure to coordinate and monitor” (p. 55), “Failure to maintain plan” (p. 57), and “Current organization inadequate” (p. 58). Non-compliance of HVB activities of written requirements in its contract with the DPED and the lack of enforcement of these requirements by the DPED are described in detail on page 67 in the audit report.

Consequently, Chapter 4’s last page offers substantial overwhelming evidence supporting the auditor’s recommendations. Here are the first two recommendations: “We recommend the following: (1) An office of tourism be established in the Department of Planning and Economic Development that is headed by a deputy director; and (2) The office of tourism make the State Functional Plan for Tourism the starting point for its tourism program” (Hawaii, 1987, p. 69).

However, these recommendations are unrealistically optimistic that a paradigm shift toward competency in managing Hawaii’s tourism program can be achieved within the existing DPED and HVB organizations. Because the incompetence is so pervasive and historical in both the DPED and HVB, such a paradigm shift may be possible to achieve only by creating a new organization to manage the state’s tourism program unique and separate from both the DPED and HVB. More recent work (i.e. the Hawaii, 1993 audit report) supports the view that the DPED and HVB have been unable to create an effective and efficient tourism-marketing program for the state.

Chapter 5: The [HVB] board of directors (21 pages). This Chapter starts by describing how unique the relationship is between the State of Hawaii and the
HVB. “The Hawaii Visitors Bureau (HVB) is both similar to – and different from – government agencies responsible for tourism promotion throughout the world. It is only one of two private, non-profit organizations promoting tourism in the United States; all other states have government-operated tourism agencies” (Hawaii, 1987, p. 71). However, Colorado voters abolished their state’s tourism board in 1993 (Bonham & Mak, 1996); such an action is one indicator that bold action can be achieved in making changes in a state’s tourism program.

The first page of the Chapter 5 summarizes the main findings:

- The Board of Directors does not play a fully meaningful role in the governance of the bureau. The bureau has not given the board the support and attention it needs to function effectively . . . ;
- The board has been unable to carry out adequately its responsibilities for policy planning, budget review, and evaluation;
- The selection processes for board directorships and leadership positions do not encourage or permit equal participation from bureau members (p. 71).

The chapter presents detailed evidence in support of these findings. The evidence is not reviewed here, but it is compelling and profound.

Chapter 6: The management of the Hawaii visitors bureau (23 pages). The main finding in Chapter 6 is stated on the first page of the chapter: “We find that: (1) There is a lack of direction and attention to several basic responsibilities in organizational management in the bureau.” Here are a few specifics to support this finding:

- We find the HVB administration does not focus sufficient attention on the organization and management of the bureau. It is not being managed as a coherent whole. The basic ingredients of sound management are lacking. The bureau has no long-range plan to give it direction, it has no formal organization chart that delineates the functions of its many units to support attainment of its goals, and its budgeting practices are hurried and slipshod. These basic deficiencies have significant ramifications for the programs it operates (Hawaii, 1987, pp. 97–98).

Chapter 6 provides detailed evidence to support these observations. The evidence includes unrealistic estimates of the private income from HVB memberships – the actual amounts are always substantially less than the budgeted amounts for 1983–1987. “The bureau compounds the problem by disregarding the actual collections for the year in budgeting for the next year – and doing this year after year” (Hawaii, 1987, p. 110).

The main recommendation of the auditor concerning management improvements of the HVB is for HVB to develop an organization plan based on a strategic plan that sets the overall direction of the program (Hawaii, 1987, p. 115). This recommendation is highly unlikely to become a reality given the overall finding
of the auditor that the HVB is “negligent in its management responsibilities.” The primary response to the main HVB related findings of the audit report support this view of reality:

While some of the findings and recommendation of the preliminary [audit] report have merit, many of the recommendations would be more appropriately addressed to a government bureaucracy rather than a private, non-profit organization. The report demonstrates an inordinate concern for developing “formal” policies and other “red tape,” and thus ignores the dynamic environment within which the HVB operates and the need for immediate, opportune response to market and media stimuli (Signed response by Walter A. Dods, Jr., Chairman of the Board, and Stanley W. Hong, President [of HVB], p. 206).

Possibly the only effective means of moving toward an effective state tourism program for Hawaii includes the complete termination of the state’s relationship with the HVB. Such a decision and action would be met with stiff resistance given the size of the HVB and the focus of its full-time staff on maintaining the status quo. Thus, Hawaii may want to consider funding the HVB with annual reductions of 25% per year from the current level of state support for four years – or reach some other solution resulting in the dissolution of its relationship with the HVB. Even though the state now could contract with a private, non-profit organization other than the HVB to operate its tourism-marketing program, the State of Hawaii would likely develop an effective tourism-marketing program by housing such a program in a separate Department of Tourism Programs.

The Hawaii Legislature should end its attempt to legislate competent behavior of the DPED and the HVB by passing additional statutes. The course of action likely to be effective includes continuing the current DPED actions of not giving the HVB direction and to continue not monitoring the activities and results of the HVB. The time has come for the State of Hawaii to create its own organization to manage the state’s tourism program.

Chapter 7: Marketing program of the Hawaii Visitors Bureau (26 pages). Chapter 7 provides substantial evidence supporting the findings offered in the first page of the chapter. “We find the following: (1) there is only limited marketing coordination among the bureau’s various departments, committees, and section offices; (2) The bureau lacks formal standards or procedures for selecting an advertising agency – its most important marketing tool and its largest cost item; (3) The bureau’s marketing program has not developed or implemented any formal monitoring and evaluation procedures or mechanisms to determine the effectiveness of its marketing activities” (p. 118).

As far as measuring the impact of the advertising and marketing efforts of the HVB, the audit report makes clear that the HVB does not follow its own Strategic Marketing Plan guidelines for creating monitoring systems.
However, despite the acknowledged importance of a monitoring/evaluation system, the HVB marketing program does not contain an integral evaluation system. Instead HVB relies on such vague ideas as the health of the industry and such haphazard methods as what people tell them (p. 138).

Even the bureau’s advertising program, the largest and the most important component of its marketing effort, is not subject to any kind of formal and ongoing monitoring or evaluation (Hawaii, 1987, p. 139). The audit report goes on to suggest several methods for evaluating advertising effectiveness – including the use of experimental designs (the Campbell & Stanley, 1963 true experiment approach, e.g. the studies reviewed by Caples, 1974 used experimental designs), as well as recall and recognition tests (these are invalid but popular methods for measuring advertising effectiveness; they are discussed in more detail below).

Given the low level of performance documented in Chapter 7, the audit report recommendations to coordinate marketing activities and to develop formal and written standards and procedures offered at the end of the chapter, while commendable, are unlikely to be accepted or adopted by the HVB. What the auditor asks for is neither “red tape” nor excessive, but rather, sound activities to achieve an effective tourism-marketing strategy.

**Chapter 8: Market research program of the Hawaii Visitors Bureau (20 pages).**

The chapter includes a benchmarking discussion of tourism research programs in the world. The main finding reported in Chapter 8: “While the Hawaii Visitors Bureau’s market research program provides the State, the bureau, and the visitor industry with valuable data, it has a number of problems which contributes to its inefficiency in certain areas, reduce its effectiveness, and limit its ability to conduct more sophisticated market research needed to improve Hawaii’s competitive edge in the world tourism-market” (Hawaii, 1987, p. 145). The audit report points out that the HVB research program is too heavily involved in statistical compilation instead of more relevant and sophisticated market research (p. 154).

The serious problems in the HVB market research extend into 1997 with the HVB commissioned study on advertising accountability study (Longwoods, 1997). This study is based on aided-recall measurement tools to measure advertising effectiveness. The results of the 1997 study indicate spectacular high performance: $75.5 million in taxes generated by visitor spending resulting from a $7.87 million advertising investment. But two points are clear from reading the scientific literature on measuring aided-recall tests: (1) they are invalid predictors of purchase; and (2) valid predictors of purchase are available (e.g. unaided top-of-mind-awareness measures and true experiments). The research design used in the HVB commissioned 1997 study is non-scientific and incorporates measurement procedures known to be invalid for estimating purchase (see
Axelrod, 1968; Caples, 1974; Haley & Case, 1979). Consequently, the findings in the HVB commissioned 1997 study are likely to be gross overstatements of Hawaii’s advertising impact on visits to the state.

Chapter 9: Visitor satisfaction and Community Relations Program (16 pages). The first finding reported in Chapter 9 is that “the visitor satisfaction and community relations program is less effective than it could be because the program lacks definition and is not integrated into the bureau’s overall marketing efforts.” The auditor recommends that the HVB develop a precise framework for the operation of its visitor satisfaction and community relations program.

Chapter 10: Personnel management 10 (pages). The main finding reported in Chapter 10 is that “The bureau has not carried out its personnel functions in a responsible and effective manner, because it does not have an ongoing personnel program” (Hawaii, 1987, p. 185). Chapter 10 presents several details to support this finding.

 Responses by the DPED and HVB
The primary response by the DPED has been discussed: the DPED disagreed with the principal recommendations in the audit report. The main response by the HVB board chairman and the president has been described: the HVB is a private non-profit organization that should not be bound by bureaucratic procedures and red tape. Other relevant points are that 80% of HVB funds came from the state; more importantly, the audit report supplies ample evidence in support of its findings, and the report’s recommendations match well with the components described earlier for effective and efficient government, tourism-marketing, programs.

Meta-evaluation of the 1987 Hawaii Management Audit Report
This report includes a useful (even though incomplete) review of relevant management and research literature pertaining to government tourism-marketing programs. The 1987 Hawaii Management Audit Report offers ample and specific evidence in support of its findings. Its recommendations are systematic in covering scanning, planning, implementing, measuring results, and administrating Hawaii’s tourism program. The recommendations are sound and their adoption would very likely increase the effectiveness and efficiency of the state’s tourism program. However, the recommendations are unlikely to be adopted, given the strong negative responses to them by the DPED and HVB. Exhibits 15 and 16
Assessing Alternative Methods of Performance Evaluation

<table>
<thead>
<tr>
<th>Theory in Use</th>
<th>Meta-evaluation</th>
<th>Rationale/additional comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Program evaluation theory used/indicated by the performance audits?</td>
<td>• Scriven-type</td>
<td>• A highly confrontational report and response by the DPD and HVB.</td>
</tr>
<tr>
<td>2. Knowledge displayed in the audit report of availability of alternative theories?</td>
<td>postpositivist</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Classification of Audit:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Comprehensive or limited in scope?</td>
<td>• Comprehensive in scope</td>
<td>Focus is on improving the quality of estimating impact of advertising</td>
</tr>
<tr>
<td>2. Focuses on evaluating which of the economic activities shown in Fig. 57 program</td>
<td>• Focused on possible cost, savings in fulfillment</td>
<td></td>
</tr>
<tr>
<td>3. Within the focus of the performance audit, provides thick descriptions of activities the done and strategies realized?</td>
<td>• No</td>
<td>• Little to no information on managing and operating activities in the report.</td>
</tr>
<tr>
<td>4. Evaluates research method used by department to measure effectiveness and efficiency of program? Audit conclusion sound?</td>
<td>• No</td>
<td>• No coverage of this topic</td>
</tr>
<tr>
<td>5. Evaluates the findings of program impact provided by the department?</td>
<td>• No</td>
<td>• No coverage</td>
</tr>
<tr>
<td>6. Reports whether or not this audit is the first, second, or third performance audit done of this department?</td>
<td>• No</td>
<td></td>
</tr>
<tr>
<td>7. If previous audits done of this department, audits reviews findings and use of previous audit reports?</td>
<td>• NA</td>
<td>• No mention of previous audit team, no scanning apparent in report</td>
</tr>
<tr>
<td>8. Reports scanning done by audit team of tourism-made by marketing program audits done by other states?</td>
<td>• No</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Assessment of Value and Impact of Audit:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Audit covers core issues related to net effectiveness/effectiveness impacts of activities and results of the program?</td>
<td>• Yes, with one issue</td>
<td>• No efficiency analyses on taxes caused by advertising; central issue not raised how many more visits occurred that would not have occurred without the advertising</td>
</tr>
<tr>
<td>2. Audit includes explicit, concrete, recommendations for change/Improvement? comprehensive</td>
<td>• Yes, very favorable</td>
<td>• How to scan and plan not covered; scientific testing advertising not covered</td>
</tr>
<tr>
<td>3. Response by program executives favorable or unfavorable toward recommendations in audit report?</td>
<td>• Unfavorable</td>
<td></td>
</tr>
<tr>
<td>4. Recommendations in report, if adopted, likely to result in substantial impacts on the quality of activities performed by the department?</td>
<td>• Yes</td>
<td></td>
</tr>
<tr>
<td>5. Central issues raised but inadequately addressed in performance audit report?</td>
<td>• No</td>
<td></td>
</tr>
<tr>
<td>6. Core issues not raised in the audit report?</td>
<td>• Yes</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Systems-Thinking and Database Marketing:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Systems thinking, display in audit report, e.g., policy mapping?</td>
<td>• Yes, but inadequately</td>
<td></td>
</tr>
<tr>
<td>2. Report covers implementation of database marketing?</td>
<td>• Yes, but inadequately</td>
<td></td>
</tr>
</tbody>
</table>

**Exhibit 15.** Results: Answers for the Issues in a Meta-evaluation of the 1989 Performance Audit of Hawaii’s Tourism-Marketing Program.
Exhibit 16. Key Issues in a Meta-evaluation of the 1989 Hawaii Performance Audit of the Division of Travel and Tourism. Note: SWOT = strengths, weaknesses, opportunity, threats; S/D = supply and demand analysis; Levels: L1 = found to be poorly done; L2 = found to be done adequately; L3 = found to be done in superior manner and/or superior performance outcome noted. Accuracy = meta-analysis evaluation of validity of auditor’s conclusion on activity/decision.

<table>
<thead>
<tr>
<th>Activity/Decision</th>
<th>Did Auditor Collect?</th>
<th>Absence Noted?</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scanning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• SWOT assessment?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Market analysis: S/D?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Trends?</td>
<td>Yes</td>
<td>No</td>
<td>L2</td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Research: customer segmenting?</td>
<td>Yes</td>
<td>No</td>
<td></td>
<td>L2</td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td><strong>Planning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Written marketing plan?</td>
<td>Yes</td>
<td>No</td>
<td>L1</td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Evidence of “what if” analyses?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Formal testing of alternatives?</td>
<td>Yes</td>
<td>No</td>
<td>L1</td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td><strong>Implementing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Milestones established?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Written monitoring guidelines</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>established/followed?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Vital competencies done in-house?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td><strong>Activity/impact assessing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Realized/planned strategy match?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Multiple dependent impact</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>measures?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Cost/benefit analysis?</td>
<td>No</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• True experiment of marketing</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>impact?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Administering</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Quality of written mission</td>
<td>Yes</td>
<td>NA</td>
<td>L3</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>statement?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Annual quantitative objectives</td>
<td>No</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>established?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Evidence of training,</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>coordination, coaching?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The course of action most likely to move the State of Hawaii’s tourism-marketing program away from its pervasive ineffectiveness and inefficiency is more a dramatic step: creating a separate Department of Tourism and ending the state’s relationship with the HVB. Given the unusually high importance of tourism for Hawaii compared to other states, and given the evidence in the 1987 audit report and more recent evidence, warrants a new state department of tourism. Some states have made dramatic changes in administering their tourism programs (e.g., Colorado in 1993, see Bonham & Mak, 1996 and our earlier discussion); consequently, enacting and implementing such a paradigm shift by Hawaii is feasible.
Based on the discussion presented above, our summary evaluation of the 1987 Hawaii audit report includes the following categories and descriptions:

- coverage of history and budget: excellent
- clarity of writing: excellent
- theory applied: a Scriven-type, post positivistic audit
- coverage of issues: comprehensive
- style: highly confrontational
- likely size of impact: little to none
- overall performance quality of report: high.

A META-EVALUATION OF A REPEAT PERFORMANCE AUDIT

The 1993 Management and Financial Audit of the Hawaii Visitors Bureau

The 1993 performance audit of the State of Hawaii’s tourism-marketing program was performed under the direction of Marion M. Higa, the State Auditor. This 32-page report includes two chapters and responses by executives of the HVB and the Department of Business, Economic Development, and Tourism (DBEPT).

Although less comprehensive, the findings and recommendations in the 1993 report are remarkably similar to the findings and recommendations in the 1987 report. However, the 1993 report notes that potential improvements in managing the tourism-marketing program were made after the 1987 audit report:

In 1990, the Legislature created the Office of Tourism in DBEDT to coordinate and plan tourism development. State funds for tourism marketing activities are channeled through this office. The office contracts with HVB and other tourism promotion programs. Currently, the Office of Tourism has separate contracts with the HVB and with each of the HVB sections on the islands of Hawaii, Kauai, and Maui. The office is responsible for monitoring HVB and performing annual reviews to ensure the effective use of state funds (Hawaii, 1993, p. 2).

expenditures and management efforts to solicit membership dues each year. Also:

Lobbying, or rent seeking, is also costly. HVB has three full-time political lobbyists, excluding its corporate officers who also actively lobby at the legislature. Moreover, HVB’s increasingly aggressive lobbying at the legislature for money has tarnished its public image and reputation. Indeed, one critique of HVB recently observed that ‘the HVB remains focused mostly not on building tourism but on building a budget’ (Bonham & Mak, 1996, p. 6; Rees, 1995, p. 5).

“To determine whether HVB has used public funds effectively, the 1993 Legislature, in House Concurrent Resolution No. 284, requested the State Auditor to conduct a management and financial audit of HVB” (Hawaii, 1993, p. 1). The 1993 audit report does provide some degree of independent confirmation of findings in the 1987 report because a different person serving as state auditor presents the findings in the 1993 audit report. However, considering that the 1993 report finds the same high level of negligence as in the 1987 report, the continued use of the HVB throughout the 1990s by the State of Hawaii is surprising. But this apparent incompetence may be explained in part by the strong tendency of most legislators not to read audit reports and by the strong, effective, lobbying efforts of the HVB.

Chapter 2: The Hawaii Visitors Bureau and its relationship with the state (12 pages). The following statement is the complete “Summary of Findings” in Chapter 2 in the 1993 audit report:

Both the Hawaii Visitors Bureau (HVB) and the Department of Business, Economic Development and Tourism (DBEDT) have fallen short in fulfilling their respective responsibilities for the State’s tourism program.

(1) The HVB board of directors has been weak and exercised little oversight over HVB.

(2) Unclear functions and underutilization of its own market research information weaken HVB’s marketing efforts. HVB has yet to resolve the status and roles of its regional offices on the mainland and its sections on the neighbor islands.

(3) HVB reports do not comply with requirements in its contract with DBEDT and do not show whether public funds are properly and effectively utilized.

(4) In the absence of strong HVB board and management leadership, DBEDT has begun to direct HVB to undertake certain programs. This conflicts with DBEDT’s responsibilities for monitoring the HVB contract (Hawaii, 1993, p. 5).

What should become clear from close reading of the 1987 and 1993 audit reports on Hawaii’s tourism-marketing programs is that the state has an unnecessary and
Assessing Alternative Methods of Performance Evaluation

A costly layer of management in these programs. Much of the effort spent in coordinating and monitoring activities between the DBEDT and the HVB could be spent more profitably elsewhere – on effectively and efficiently managing government tourism-marketing programs. Even without the continuing poor tourism-marketing performance of the HVB, the State of Hawaii would benefit by ending its unique arrangement of contracting out for vital competencies that need to be mastered within the Office of Tourism.

The recommendations in Chapter 2 include reducing the number of directors on the HVB board, providing written guidelines for committee members, and developing a (written) strategic plan. The recommendations include the DBEDT “submitting annual reports to the Legislature that contain the information requested by the Legislature on tourism promotion programs and their effectiveness” (Hawaii, 1993, p. 17).

The HVB and DBEDT Responses and the Auditor’s Response to the Responses

The HVB president commented that many of the audit report’s recommendations reflect inordinate concern for developing “formal” practices. The auditors’ response, “We note that a certain minimum level of written procedures and instructions is required for any organization to operate effectively.

The director of the DBEDT did not respond directly to recommendations. “Permit me to offer the following general comments, rather than a point-by-point rebuttal,” stated Mufi Hanneman, Director of the DBEDT, by letter on December 27, 1993 (Hawaii, 1993, p. 28). The audit report criticizes the DBEDT for having the HVB undertake tourism initiatives developed by the DBEDT. The director’s response defends the department’s authority to undertake tourism initiatives regardless of whether they are planned or not. The auditor’s response, “We believe, however, that the department should undertake these initiatives in its own name and not under that of HVB” (Hawaii, 1993, p. 21).

Meta-evaluation of the 1993 Hawaii Management and Financial Audit Report on the HVB and Its Relationship with the State

The 1993 report supports the main findings strongly using detailed evidence. However, the range of topics covered in the report is disappointing. Whereas the 1987 Hawaiian audit report is comprehensive in coverage, the 1993 audit report is limited in coverage. Unlike the 1987 audit report, the 1993 report does not include the attempt to identify best practices in government tourism-marketing programs. The 1993 report does not include detailed information on the scanning and planning activities of the HVB and the DBEDT. The 1993 audit report does not
### Exhibit 17. Results: Answers for the Issues in a Meta-evaluation of the 1993 Performance Audit of Hawaii’s Tourism-Marketing Program.

<table>
<thead>
<tr>
<th>Theory in Use</th>
<th>Meta-evaluation</th>
<th>Rationale/additional comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Program evaluation theory used/indicated by the performance audits?</td>
<td>• Accounting theory of program evaluation</td>
<td>• A mildly confrontational report and response by the tourism division.</td>
</tr>
<tr>
<td>2. Knowledge displayed in the audit report of availability of alternative theories?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classification of Audit:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Comprehensive or limited in scope?</td>
<td>• Limited in scope</td>
<td></td>
</tr>
<tr>
<td>2. Focused on evaluating which of the department activities shows in Fig. 5?</td>
<td>Focused on lack of coordination and DBEDT directing IVB</td>
<td></td>
</tr>
<tr>
<td>3. Within the focus of the performance audit, provides thick descriptions of activities done and strategies realized?</td>
<td>• No</td>
<td></td>
</tr>
<tr>
<td>4. Evaluates research method used by department to measure effectiveness and efficiency of program? Audit conclusion sound?</td>
<td>• No</td>
<td>No coverage of this topic</td>
</tr>
<tr>
<td>5. Evaluates the findings of program impact provided by the department?</td>
<td>• No</td>
<td>No coverage</td>
</tr>
<tr>
<td>6. Reports whether or not this audit is the first, second, or third performance audit done of this department?</td>
<td>• Yes</td>
<td></td>
</tr>
<tr>
<td>7. If previous audits done of this department, reviews findings and use of previous audit reports?</td>
<td>• Yes</td>
<td></td>
</tr>
<tr>
<td>8. Reports scanning done by audit team of tourism-maintained marketing program audits done by other states?</td>
<td>• No, none apparently done</td>
<td>• No claims of scanning audit team; no such scanning apparent in report</td>
</tr>
</tbody>
</table>

### Assessment of Value and Impact of Audit:
1. Audit covers core issues related to net effectiveness/efficiency impacts of activities and results of the program? • No • No efficiency analyses on taxes caused by advertising; central issue not raised how many more visits occurred that would have occurred without the advertising
2. Audit includes explicit, concrete, not recommendations for change/improvement? very limited • Yes, but unfavorable
3. Response by program executives favorable or unfavorable toward recommendations in audit report? • Unfavorable
4. Recommendations in report, if adopted, likely to result in substantial impacts on the quality of activities performed by the department? • No • How to scan and plan not covered; scientific testing advertising not covered
5. Central issues raised but inadequately addressed in performance audit report? • Yes • What are the net tax
6. Core issues not raised in the audit report? • Yes

### Systems Thinking and Database Marketing:
1. System thinking display in audit report, e.g. policy mapping? • No
2. Report covers implementation of database marketing? • No
include a financial audit report, even though “financial audit” appears in the title of the report. The 1993 audit report fails to examine the effectiveness and efficiency of the HVB’s tourism-marketing programs. Other than reporting that the “HVB reports do not comply with requirements in its contract with DBEDT and do not show whether public funds are properly and effectively utilized (Hawaii, 1993, p. 5);’ the 1993 audit report fails to report HVB’s own attempts, if any, to measure the effectiveness and efficiency of its tourism-marketing programs.

In brief, the 1993 audit report is substantially lower in quality compared with the 1987 audit report. Exhibits 17 and 18 summarize our meta-evaluation of the 1993 audit report.

<table>
<thead>
<tr>
<th>Activity/Decision</th>
<th>Did Auditor Collect?</th>
<th>Absence Noted?</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>Accuracy?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scanning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• SWOT assessment?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Market analysis: S/D?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Trends?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Research: customer segmenting?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Written marketing plan?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Evidence of “what if” analyses?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Formal testing of alternatives?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Milestones established?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Written monitoring guidelines</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>established/followed?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Vital competencies done in-house?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity/Impact assessing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Realized/planned strategy match?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Multiple dependent impact measures?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Cost/benefit analysis?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• True experiment of marketing impact?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administering</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Quality of written mission statement?</td>
<td>Yes</td>
<td>NA</td>
<td>L3</td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Annual quantitative objectives established?</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Evidence of training, coordination, coaching?</td>
<td>Yes</td>
<td>Yes</td>
<td>L1</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
</tbody>
</table>

*Exhibit 18.* Key Issues in a Meta-evaluation of the 1993 Hawaii Performance Audit of the Division of Travel and Tourism. *Note:* SWOT = strengths, weaknesses, opportunity, threats; S/D = supply and demand analysis; Levels: L1 = found to be poorly done; L2 = found to be done adequately; L3 = found to be done in superior manner and/or superior performance outcome noted. Accuracy = meta-analysis evaluation of validity of auditor’s conclusion on activity/decision.
The following categories and descriptions are a further summary to the 1993 Hawaii management audit report of the HVB:

- coverage of history and budget: incomplete
- clarity of writing: good
- theory applied: a Scriven-type, post positivistic audit
- coverage of issues: limited
- style: confrontational
- likely size of impact: little to none

A META-EVALUATION OF A PERFORMANCE AUDIT OF A NATIONAL TOURISM-MARKETING ORGANIZATION

The 1993 Australian Audit Report of the Australian Tourist Commission’s Marketing Impact

The 1993 Australian audit report, Evaluation of the Australian Tourist Commission’s Marketing Impact, is 73 pages followed by more than 100 pages of appendices. The Australian audit report is the only one of the seven reviewed that was “compiled” by three organizations, rather than one government audit organization. The three compilers are The Department of Finance (DOF); Department of Arts, Sport, the Environment, Tourism and Territories (DASET); and the Australian Tourist Commission (ATC).

The Australian audit report has an executive summary and eight chapters. The report is the only example of a partnering-type performance audit among the seven audits examined in this meta-evaluation (see Wholey, 1983). The main conclusion of the audit is a positive one: “The report concludes that ATC has been effective in meeting its marketing objectives as set out in its Act” (Australia, 1993, p. 1).

The other main findings are positive as well:

- ATC provides important value added and, therefore, contributes to tourism receipts which benefit tourism industries generally and the community as a whole;
- The impact of ATC marketing, though difficult to quantify, is positive;
- Continued Government funding of international tourism promotion is justified, in conjunction with – where possible – increased private sector contribution to joint activities (Australia, 1993, p. 1).
The audit was intended to be the “first comprehensive evaluation of the Australian Tourist Commission’s (ATC) marketing and promotional activities.” Unfortunately, the report does not evaluate many of the activities of the ATC directly. For example, the audit report does not report whether or not the ATC prepares a written marketing plan. Consequently, the report does not address the quality of planning done by the ATC before the audit.

The audit report fails to describe how the ATC is organized; the size of the ATC; and its budget and disbursements. How did the ATC evaluate its impact before the auditing team did its study? The report does not address these issues. Detailed descriptions and rationales of the specifics of Australia’s marketing and advertising campaigns are not provided in the 1993 Australian audit report. The report provides a conglomeration of examples of numbers of consumers reached by Australian’s tourism advertising for 1989, 1990, and other years (no annual breakouts are provided). The report describes conversion studies without comparing the relative performances of different media vehicles, or different advertising-segment campaigns – the most appropriate use for conversion studies (see Woodside & Motes, 1981).

What is the tax profit from the Australian tourism-marketing program net of expenditures for the program? The report does not provide estimates in response to this issue. Although the 1993 Australian report is attractive in appearance, it fails to provide a cost-benefit analysis of the Australian tourism-marketing program.

The auditors writing the report have a naïve understanding of “control experiment” (p. 52). “One of the problems throughout this evaluation has been the difficulty in conducting a controlled experiment by manipulating the variables in the decision-making process. All these variables, such as disposable income, airfares and exchange rates are beyond the control of researchers” (p. 52). True-experiment designs for examining the effects of advertising include manipulating one or more advertising variables – using one or more treatment groups and a control group. See the discussion of true experiments presented earlier in this meta-evaluation for a detailed description of true experiments in relation to measuring effectiveness of tourism-marketing and advertising programs.

The audit reports discusses the use of aided recall for measuring the impact of Australian’s tourism advertising. For example, the question is raised, “have you seen advertising for visiting Australia?” Among respondents in a study in Japan, 19% reporting Australia in first-mention responses to the question on intent to visit Australia in the next three years versus 10% of respondents in the “not seen” first-mention category (p. 53). Given the inherent biases in aided-recall questions, given the lack of validity of the method in measuring advertising effectiveness in the literature, and given the long-time availability and wide applications of scientific methods to measure advertising effectiveness (see Banks, 1965;

<table>
<thead>
<tr>
<th>Theory in Use</th>
<th>Meta-evaluation</th>
<th>Rationale/additional comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Program evaluation theory used/dominated by the performance audits departments</td>
<td>• No specific grounding</td>
<td>• Three government participated in preparing the audit</td>
</tr>
<tr>
<td>2. Knowledge displayed in the audit report of availability of alternative theories to be</td>
<td>• Limited in scope</td>
<td>• Claim made early in report comprehensive</td>
</tr>
<tr>
<td>3. Within the focus of the performance audit, provides thick descriptions of activities done and strategies realized?</td>
<td>• No</td>
<td></td>
</tr>
<tr>
<td>4. Evaluates research method used by department to measure effectiveness and valid efficiency of program? Audit conclusion sound?</td>
<td>• No</td>
<td>• Accepts methods used as</td>
</tr>
<tr>
<td>5. Evaluates the findings of program impact provided by the department?</td>
<td>• No</td>
<td>• No coverage of this topic</td>
</tr>
<tr>
<td>6. Reports whether or not this audit is the first, second, or third performance audit done of this department?</td>
<td>• Yes, the first audit</td>
<td></td>
</tr>
<tr>
<td>7. If previous audits done of this department, reviews findings and use of previous audit reports?</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>8. Reports scanning done by audit team of tourism-marketing program audits done by other states?</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>

| Assessment of Value and Impact of Audit: |
| 1. Audit covers core issues related to net effectiveness/efficiency impacts of activities and results of the program? | • To a limited degree | • No efficiency analyses on taxes caused by advertising; central issues not raised how many more visits occurred that would not have occurred without the advertising |
| 2. Audit includes explicit, concrete, not recommendations for change/improvement? very limited | • Yes, but | |
| 3. Response by program executives favorable or unfavorable toward recommendations in audit report? | • Favorable; ATC co-author of report | |
| 4. Recommendations in report, if adopted, likely to result in substantial impacts on the quality of activities performed by the department? | • Few specific recommendations in report | |
| 5. Central issues raised but inadequately addressed in performance audit report? | • Yes | |
| 6. Core issues not raised in the audit report? | • Yes, several | • How to scan and plan not covered; scientific testing advertising not covered |

| Systems Thinking and Database Marketing: |
| 1. Systems thinking displayed in audit report, e.g. policy mapping? | • No | • What are the net revenues and costs/benefits of tourism-marketing program? |
| 2. Report covers implementation of database marketing? | • No | |
| Database | | |

644 ARCH G. WOODSIDE AND MARCIA Y. SAKAI
Assessing Alternative Methods of Performance Evaluation

Caples, 1974; Ramond, 1966), the use of aided recall measures by the ATC is surprising. As noted above, this use indicates a lack of knowledge of and training in scientific measurement of program impact (see Banks, 1965; Campbell, 1969; Woodside, 1990).


The report presents no history on the creation, budgets, and performance history of the ATC. The report does not describe specific annual marketing plans. The report

<table>
<thead>
<tr>
<th>Activity/Decision</th>
<th>Did Auditor Collect?</th>
<th>Absence Noted?</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>Accuracy?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scanning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• SWOT assessment?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Market analysis: S/D?</td>
<td>Yes</td>
<td>Present</td>
<td>L3</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Trends?</td>
<td>Yes</td>
<td>Present</td>
<td>L3</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>• Research: customer segmenting?</td>
<td>Yes</td>
<td>Present</td>
<td>L3</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
<tr>
<td>Planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Written marketing plan?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Evidence of “what if” analyses?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Formal testing of alternatives?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Implementing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Milestones established?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Written monitoring guidelines established/followed?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Vital competencies done in-house?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Activity/Impact Assessing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Realized/predicted strategy match?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Multiple dependent impact measures?</td>
<td>Yes</td>
<td>Present</td>
<td>L3</td>
<td></td>
<td></td>
<td>Inaccurate</td>
</tr>
<tr>
<td>• Cost/benefit analysis?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• True experiment of marketing impact?</td>
<td>Yes</td>
<td>Yes</td>
<td>L3</td>
<td></td>
<td></td>
<td>Inaccurate</td>
</tr>
<tr>
<td>Administering</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Quality of written mission statement?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Annual quantitative objectives established?</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>• Evidence of training, coordination, coaching?</td>
<td>Yes</td>
<td>Present</td>
<td>L2</td>
<td></td>
<td></td>
<td>Accurate</td>
</tr>
</tbody>
</table>

Exhibit 20. Key Issues in a Meta-evaluation of the 1993 Australian Performance Audit Report. Note: SWOT = strengths, weaknesses, opportunity, threats; S/D = supply and demand analysis; Levels: L1 = found to be poorly done; L2 = found to be done adequately; L3 = found to be done in superior manner and/or superior performance outcome noted. Accuracy = meta-analysis evaluation of validity of auditor’s conclusion on activity/decision.
does not describe how scanning activities are used to create marketing plans. The report does not describe the present organization and responsibilities of persons in management positions. How did the results of prior advertising-impact studies influence current decisions of the ATC? What are the cost-benefit estimates for specific advertising campaigns in specific countries? Does the ATC have overseas offices? If so, how are the activities of these overseas offices evaluated for their effectiveness and efficiency? The 1993 Australian audit report does not address these issues.

Exhibits 19 and 20 provide summaries of our meta-evaluation on the Australian 1993 audit report. Based on the discussion presented above, our summary evaluation of the 1993 Australian audit report includes the following categories and descriptions:

- coverage of history and budget: poor
- clarity of writing: poor
- theory applied: no specific grounding
- coverage of issues: limited to minor issues
- style: partnering
- likely size of impact: moderate

CONCLUSIONS AND IMPLICATIONS FOR META-EVALUATION RESEARCH ON PERFORMANCE AUDITS OF TOURISM-MARKETING PROGRAMS

For the most part the government performance audits reviewed indicate that comprehensive audits were not performed of either the activities or impacts of their governments’ tourism-marketing program. With the exception of the 1987 Hawaii audit report, three vital questions in tourism-marketing stand out as not being addressed adequately by either the government agency responsible for the tourism-marketing program or by the auditors:

- What are the profits in net taxes resulting from the government’s tourism-marketing program?
- Are the activities required (see Fig. 5) for achieving an effective and efficient tourism-marketing program being performed well by the government agency responsible for administering the program?
- What are the environmental and community impacts of the tourism-marketing program and the entire state or country tourism program?

The audit reports reviewed for this meta-evaluation report do not indicate training in or knowledge of the field of program evaluation. References to the leading
theories of program evaluation do not form the foundation of any of the audit reports reviewed. Although the Australian (1993) audit report claims that the audit team completed reviews of state-of-the-art evaluation techniques, the report does not include evidence of even rudimentary awareness of the journals and literature of the field of program evaluation. A statement favored for theoretically grounding the performance audits reviewed is this one: “Our work was conducted . . . in accordance with generally accepted government auditing standards” (Hawaii, 1993, p. 3). If accurate, then the generally accepted government auditing standards are out-of-date, limited in scope, and lack requirements for cost-benefit analyses.

As found in this meta-evaluation, the depth of performance audits is a continuum from very limited to very comprehensive. While limited performance audits may be necessary for some specific situations, depth of performance audits is not independent from audit quality. Everything else being equal, the performance audit covering and reporting detailed findings of all managing topics (i.e. scanning, planning, implementing, measuring results, and administrating) will be more useful and higher in quality than the audit limited in scope. The worry is that limited-focused audits tend to result in the non-audit-audit as illustrated by the 1995 Tennessee audit. Thus, although possibly more efficient, the limited-focused audit is more often less effective than the comprehensive audit.

Little evidence was found in this meta-evaluation that audit findings and recommendations have immediate impact on changing behavior. At best, they appear to provide some enlightenment to alternative views of reality. However, some changes other than enlightenment do result from audits; the creation of the Office of Tourism in Hawaii in 1990 is likely due, in part, to the attention paid by the Hawaiian Legislature to the results of the 1987 audit. Including a detailed description in the 1993 Hawaii audit of what happened in 1987 through 1990 by government agencies using the 1987 performance audit report would have improved the quality of the 1993 report. Ultimately, increasing enlightenment and perseverance among a few persons (e.g. Legislature members and tourism-marketing program directors) sometimes results in improved sensemaking and effective and efficient performance.

However, both audit teams and legislators might consider creating organizational mechanisms to evaluate recent performance audits (a self-meta-evaluation) and to revise and revising and implementing a few recommendations in the audit. These steps are missing now from performance audit practice. We suggest creating an audit-report-utilization “working team” of three parties; the team would include three to seven persons chosen from the legislature, the tourism-marketing program, and the audit office. We have no time for such an additional audit-utilization step – is the likely cry by auditors, program directors, and legislators. However, the alternative is continuing to focus on doing
more audits that no one reads as opposed to shifting to doing fewer audits that get used.

Future research on the meta-evaluations of performance audits of government tourism-marketing programs will benefit from direct interviews with audit team members doing the fieldwork for the program evaluations. In addition, meta-evaluation research studies are needed to estimate the relative impacts of alternative advertising campaigns.

Although single-group conversion-study designs do not address the issue of how many visitors exposed to the advertising and who also visited would have come without being exposed, such no-control group conversion studies are valuable for examining the relative impacts of competing advertising messages and advertising media. Meta-analyses (see Hunter & Schmidt, 1990) are needed for estimating conversion rates across studies; and for testing different moderator variables, for example, the effects of advertising themes, different direct-response campaigns, different customer segments, and different years of program operation on conversion rates. Detailed reporting of conversion rates for different media buys, markets, ads, and customer segments would provide valuable data for such meta-analysis studies (e.g. see Woodside & Motes, 1981).

Is the performance audit agency improving in its auditing activities and the impacts of its reports? Is the auditing agency up-to-date in grounding its research and reporting in the literature on program evaluation? We need additional meta-evaluation studies to answer these issues.

Because state agencies and all strategic business units are responsible for self-assessments of activities and impact, performance auditors need to accept greater responsibility for training the agency being audited on the methods of program evaluation – including the relevant literature of program evaluation. The auditors need to display the expertise necessary for sound performance audits.

BUILDING A NEW PARADIGM: FIVE GOLDEN RULES FOR PERFORMANCE AUDITING OF TOURISM-MARKETING PROGRAMS

This section describes five golden rules for performance auditing of tourism-marketing programs. These golden rules are intended to be a composite practical model of performance auditing of government tourism-marketing programs.

*Golden Rule 1: Embrace On-going Formal Training in Program Evaluation*

On-going formal training in program evaluation theory and method is necessary for both auditors and executives of tourism-marketing programs. A major finding
of the present meta-evaluation is the absence of references and understanding of program evaluation knowledge and literature by the program evaluators and tourism-marketing executives.

Some of the audits we have reviewed make limited reference that standard accounting audit procedures were applied. None of the audits provide detailed descriptions of the coverage in accounting audits. More importantly, the reports offer no theoretical ground of what constitutes sound marketing practice and performance outcomes. The performance audits reviewed do not explicitly state and build on a framework (i.e. a mental model or set of assumptions) of what constitutes sound tourism-marketing practice or impact.

The seven audits include no references to the scientific literature of program evaluation. Consequently, useful theoretical and practical foundations of program evaluation knowledge (see Shadish, Cook & Leviton, 1991) are absent from these reports.

Consequently, most of the reports do not address the epistemology for doing sound performance audits and implementing sound marketing strategies. The 1985 Minnesota audit and the 1993 Australian report come closest to explicating an epistemology of evaluating tourism-marketing programs. Both the Minnesota and the Australian reports raise a fundamental issue in epistemology: how to build into our marketing programs the comparisons of outcomes of planned marketing action vs. planned non-action. However, neither these two reports nor the others go very far in identifying the necessary improvements in tourism-marketing strategies and program evaluation methods – the reports do not advocate split-run testing (i.e. true experiments in marketing, see Banks, 1965; Woodside, 1990).

Embracing formal training in program evaluation should include reading seminal works in program evaluation (e.g. Shadish et al., 1991) and scientific evaluating methods of marketing actions (e.g. Banks, 1965; Caples, 1974). Reading meta-evaluations of tourism-marketing evaluations will be useful as well. Acquiring deep knowledge about the literature on program evaluation leads to building sound explicit models of program evaluation beyond the naïve approaches in use in unwritten implicit models.

**Golden Rule 2: Audit both Program Activities (Implemented Strategies) and Impacts on Planned Objectives**

For several reasons measuring outcomes only is not good enough. Unfortunately, outcome measurement of tourism-marketing programs is done without planned control-group comparisons. The relatively larger sizes of environmental factors may make systematic program influences difficult to uncover. Sampling and non-sampling errors may occur in measuring outcomes that degrade the information and
increase the level of “noise” in the outcome data. Finally, using multiple indicators of performance of actions taken provides convincing evidence of impact.

Program performance measurement needs to be inclusive of both process and outcome. The core issue to raise is whether or not the government’s tourism-marketing program generated visits that would not have occurred without the program? The performance audit should raise another core issue: what actions were implemented in administering, scanning environments, planning, implementing, and self-assessing of performance that caused the impact?

Creating an explicit, written, plan of activities helps to improve thinking and action. Writing an annual marketing plan forces executives to make explicit their implicit models of how tourism-marketing strategy is done effectively. Writing a marketing plan requires reflection and deepens sensemaking (see Weick, 1995). Written plans include important implementation details often overlooked in verbal plans. Detailed mapping of the actions and feedback loops lead to dramatic improvements in systems thinking. Such mapping provides knowledge that can be gained in preparing detailed written plans.

However, only the audits for Hawaii’s tourism-marketing programs raised issues related to the preparation and use of a marketing plan. None of the seven audits raised the fundamental issues of:

- What constitutes a high-quality marketing plan?
- What is the dominant mental model displayed in the marketing strategy actually implemented in the current tourism-marketing program?
- Should the present dominant model be replaced with a new paradigm? If yes, what are the details in this new mental model?

Detailed descriptions of what was done in the planning and implementing of the tourism-marketing program should be an integral part of the performance audit. Whether or not a written annual marketing plan was prepared should be included in such detailed descriptions. If a plan was prepared, the audit should address the quality of the plan. The detailed description should include details on what actions were implemented and the degrees of fit of the implemented strategy with the planned strategy. Also, the audit should ask the program participants in implementing the strategy to evaluate their actions and actions not taken that should have been taken.

While participants are asked to generate a lot of data in program evaluations, rarely are they directly asked to evaluate the program, to judge its adequacy, to advise on its continuance, discontinuance, dissemination or modification. Rather than evaluating programs, participants are usually asked about themselves and their own adequacy. We are thus wasting a lot of well-founded opinions (Campbell, 1978 in Overman, 1988, p. 374).
The seven audits reviewed do not include the participants’ views on the quality of the implemented strategies as integral parts of the audit reports. With the exception of the Australian audit report, senior program executives provided comments only on the accuracy of the audit and only after reading the written audit report. This procedure is not the same as seeking the views on the quality of the strategy implemented by participants as integral to the auditing process. Also, the detailed descriptions need to include the views on program performance of section managers and service workers. Applying the old saw, “God is in the details,” requires learning the details about implementation unlikely known senior program executives – as well as details about how implementation could have been done better.

Qualitative assessment of activities serves to confirm and extend evaluation of quantitative results. Designing both evaluation methods into performance audits is likely to confirm the major findings on their shared dimensions. Adopting this golden rule should include embracing the following view by Campbell. If disconfirmation of findings on shared dimensions is found, “we should consider the possibility that the quantitative procedures are in error. If I will concede this, why would I be reluctant to see the qualitative procedures used without the quantitative? It is because I believe that the quantitative, when based on firm and examined qualitative knowing, can validly go beyond the qualitative, and can produce subtleties that the qualitative would have missed” (Campbell, 1978 in Overman, 1988, p. 375).

Golden Rule 3: Embrace a New Multi-theory Based Paradigm for Performance Audits of Tourism-Marketing Programs that Includes Stakeholder Participation

Among the seven audits reviewed, the dominant paradigm displayed in conducting the performance audits of tourism-marketing programs is confrontational. The lone exception is the Australian 1993 audit – a performance audit planned and implemented by executives of three agencies including executives of the agency being audited.

The case for adopting a new paradigm that includes stakeholder participation in designing the audit includes several dimensions. First, Stake (1980) provides a well-developed theory with rationales for adopting what he refers to as a “responsive evaluation” (as we described in Section 2). Second, Campbell expressed one of the core rationales of Stake: stakeholders are likely to provide unique and valuable information on implementation problems and opportunities – as well as creative insights on improving the planning of next year’s
tourism-marketing programs. Third, earlier sections provide strong evidence in this meta-evaluation that supports the following conclusions: (1) tourism-marketing program executives fight against making the changes and adopting improvements recommended by auditors; and (2) the executives win these fights. The result is the implementation of only minor changes. Witness the battle over a number of years between the Hawaii audit office vs. the HVB and the leaders of the Hawaii’s tourism-marketing program. Even with the heavy weight of the Hawaii legislative branch supporting the conclusions and recommendations made by the auditor, no major improvements in the activities, performance outcomes, or measurement methods has occurred in the Hawaii tourism-marketing program.

Embracing Golden Rule 3 implies partnership, not audit leadership, by the executives of the tourism-marketing program. The objective of such an audit becomes less on passing judgment on the lack of planning and poor performance of individuals and more on documenting the strategy paradigm implemented in the current tourism-marketing program – and comparing the strategy implemented with alternative, potentially more effective and efficient, strategies. See Golden Rule 4 below.

Golden Rule 4: Transform Government Tourism-Marketing Strategy from Transactional to Relationship Marketing

Figure 10 summarizes the dominant strategy paradigm in use in government tourism-marketing programs. Over 50% of the total marketing funds budgeted is spent in steps one and three in Fig. 10. In step one, the placement of advertising in scheduled media, mostly in newspapers, magazines, television, usually represents half or more of the total tourism-marketing budgets.

Note that Fig. 10 includes one contact only with inquirers responding to the offer of free literature. After fulfillment of the requests, the names, addresses, and related information (e.g. media vehicle) are discarded. This strategy includes no efforts or budget for development of an on-going relationship with the inquiring prospective visitors. The strategy includes making no additional contacts after mailing the free “linkage advertising” (Rapp & Collins, 1988) literature. Almost all government tourism-marketing programs do not practice database marketing, even though some departments provide inquiry names to industry trade associations.

Transforming tourism-marketing programs from the currently dominant transactional paradigm to a relationship-marketing paradigm is justified on several grounds. First, inquirers are persons who have identified themselves as having a high interest in buying the product (i.e. visiting the destination). Even if most do
Assessing Alternative Methods of Performance Evaluation

Fig. 10. The Dominant Two-Step Marketing Strategy Now Being Used in Government Tourism-Marketing Programs. Note: The strategy is "two-steps" because the marketer makes two attempts to attract visitors: (1) image advertising followed by; (2) the linkage advertising of the free visitor's guide. The typical conversion study following the implementation of such strategies are not valid for measuring the share of destination visitors exposed to the image and/or linkage advertising who would have not otherwise have visited the destination.
not buy during the current buying season, the strong potential exists that some may be converted into product buyers (i.e., visitors) if contacted more than once over the next two buying seasons. Tourism-marketing executives can test this proposition scientifically— with test and control groups—and then, estimating the profitability of multiple-contacts customers is possible.

Second, the current dominant marketing strategy includes substantial image and linkage advertising to identify inquirers by names, addresses, and inquiry sources, as well as to fulfill their requests for information. Estimating the long-term net benefits from these expenditures is impossible if no database is created and used that contains detailed information by individual customer files.

Third, because of the advances in computer technology and reductions in software costs, the costs of data handling and storage has declined dramatically over the past three decades. In 1993, Bulkeley made this point dramatically, ”High-end pc’s with two-gigabit hard drives – 20 times faster than the 100 megabyte drives most home users buy now – can hold several million customer names on hardware that costs about $10,000 now. The software to manage such data starts as low as $15,000” (B6).

Fourth, the strategy paradigm for database or “maximarketing” is well developed now (see Rapp & Collins, 1988; Woodside, 1996). Most large private enterprises have transformed their marketing strategies to database marketing, for example, Sears, Nordstrom, General Motors, Papa John’s Pizza, Pfizer, Schwab, and Delta Air Lines. Consequently, transforming a government tourism-marketing program to database marketing does not have to be a shot-in-the-dark: many firms in many industries completed the transformation years ago.

For some of the media vehicle sources, the average trip expenditure made by inquirers who do visit the advertised destination is greater than $500 for the one trip to the destination— for some media vehicles the average expenditures per converted inquiry party is greater than $1,000. A 1987 inquiry study for Louisiana notes that the average expenditures in Louisiana among inquirers from an ad placed in Travel and Leisure who converted into visitors was $1,208 in 1987 dollars (see Woodside, 1996, p. 133). Consequently, developing a relationship-marketing program aimed at identifying and nurturing special long-term relationships with such valuable customers makes more sense than continuing the use of one-time transactional marketing.

Figure 11 includes some details of the start-up of one possible tourism relationship-marketing program. Note that in steps 6 and 8 in Fig. 11 includes making multiple contacts with first inquirers. The destination marketer provides these inquirers with the option of ending all future contacts. A core proposition in such database marketing is that “mining” inquiries by learning the individual travel interests of inquirers and designing specialized marketing appeals to
Assessing Alternative Methods of Performance Evaluation

Fig. 11. Steps in Building Toward a Relationship-Marketing Program. Note: The strategy includes multiple contacts with inquirers by creating a database marketing program. Prospects are segmented into interest groups and purchase response categories for more individualized future mailings. Prospects who agree to participate in the long-term relationship but do not respond after 3–5 contacts with a purchase (i.e., a visit to the destination) may be dropped from the program or may be asked to pay a premium for their continued participation.
special interest groups (SIGs) is highly profitable – more profitable than making one response only to inquirers. The literature on database supports this core proposition (e.g. Rapp & Collins, 1988).

Given a set budget constraint, the shift from transactional marketing to database relationship marketing does result in shifting some share of a scheduled media budget to nurturing one-on-one relationships. Even though advertising agencies have fought such shifts in expenditures (see Cappo, 1992; Woodside, 1996), most firms in most industries made the transformation without overwhelming reductions in scheduled media budgets.

Also, the tourism-marketing program managers can expect a decline in first-time inquiries with a transformation to relationship marketing. The focus on the value of individual customers for the destination becomes much greater with relationship marketing compared to transactional marketing.

Transforming the marketing program to relationship marketing requires providing evidence of its greater profitability compared to the traditional two-step only marketing program. Conversion studies can be transformed into treatment group vs. control group experiments to examine the value of relationship marketing programs. Given that individual names, addresses, and relationships with the destination are known, random samples can be created to receive special marketing offers or no offers at all. Such research designs are the basis for scientific testing of the impact of marketing programs on sales and profits (see Banks, 1965).

Golden Rule 5: Do Split-Run Tests of Specific Marketing and Advertising Actions

Given that the primary interest regarding tourism advertising and marketing of government legislative members is whether or not the advertising caused visits that would not otherwise have occurred, the following conclusion may be surprising. No scientific tests of effectiveness have been done on the impact of government tourism advertising on influencing visits. Scientific testing of a treatment’s influence on a dependent variable requires direct comparisons of a treatment group with a comparable control group.

The lack of scientific testing of advertising effectiveness is even more surprising for the following additional reasons. First, the details on how to perform such tests methods are widely known (see Caples, 1974; Faulkner, 1997; Woodside, 1996). Second, most government tourism-marketing departments spend substantial funds each year for measuring the impact of advertising on influencing visits. Third, each year legislators raise the primary issue regarding government tourism advertising: did the advertising generate visits that would not have occurred otherwise?
We can speculate on at least four reasons for our lack of scientific data on tourism advertising’s influence on visits. First, tourism-marketing executives, tourism-advertising researchers, elected representatives, and performance auditors may lack the knowledge of and training in how to estimate advertising-to-sales causation scientifically – even though such knowledge and training is readily available (see Banks, 1965; Caples, 1974; Woodside, 1990).

Second, non-scientific testing, such as frequently used conversion studies and aided recall surveys, appears to provide useful information for answering the core issue (e.g. Longwoods International, 1997; Woodside & Soni, 1990). However, the available conversion studies and aided recall surveys do not meet the necessary requirement of scientific testing: such studies include no planned comparisons of persons exposed to the advertising vs. comparable persons not exposed to the advertising. Thus, the typical conversion study report usually ends with the misleading implication that advertising resulted in visits by 40 or 50% of the inquirers generated from the tourism-advertising program who visited the destination after receiving the free literature they requested.

Asking respondents to indicate whether or not the scheduled media advertising and free literature influenced their decision to visit is not a scientific measurement of cause-and-effect relationships. Self-reports of causal attribution are often inaccurate and distorted by psychological processes (see Plous, 1993).

Third, in a number of scientific tests, detecting the impact of advertising on sales is not possible but the impact is very small. Or worse: the finding is reached that advertising decreases sales – such a result has been found to occur in at least one scientific test (see Ackoff & Emshoff, 1975). Consequently, why take the chance of doing a scientific study when the results may be unfavorable?

Thus, fear of unwanted results may be a fourth reason why scientific testing of advertising’s impact on visits has not occurred. Scientific test results may result in a poor performance evaluation of the advertising program, and by association, poor performance by the advertising agency and the tourism-marketing executives. By comparison, the widely used single-group conversion study and aided-recall surveys result in feel-good findings. Such studies nearly always result in favorable findings that advertising influences visits and huge expenditures by tourists, for example, 50% inquiry-to-visits conversion and $200 expenditures by tourists for every $1 of scheduled media advertising.

To accomplish the paradigm shift in thinking necessary to start using scientific tests of tourism advertising effectiveness, we suggest taking the following steps. First, one big scientific test should not be run. Designing and implementing small scale tests decreases the risk that the results of any one study provide inaccurate findings regarding the effectiveness of advertising generating visits. Applying
the counsel on quantitative program evaluation offered by Campbell (1978 in Overman, 1988, p. 372) is wise:

Quantitative experimental studies involve so many judgmental decisions as to mode of implementation, choice and wording of measures, assembly of data in analysis, etc., they too should be done in replicate. Our big evaluations should be split into two or more parts and independently implemented. When the results agree, the decision implications are clear. When they disagree, we are properly warned about the limited generality of the findings.

Second, a research design using advertising free standing inserts (FSIs) in magazines and newspapers offer a low cost application for scientific testing. For example, we offer the following suggested steps: select 2,000 subscribers in each of three magazines and three newspapers systematically. For the study, create a full-page advertisement as a FSI. Assign half of the subscribers (1,000) in each of the six media groups randomly to receive the FSI in their regular issue of the media vehicle – the other half (1,000) receives no FSI in the same regular issue of the media vehicle. Thus, for each media vehicle, two scientifically comparable groups are created, with the treatment group exposed to the tourism advertising and the control group not exposed to the tourism advertising. Because of the random assignment of households to the two groups, we can eliminate sources of influence other than the advertising as generating found differences in visits. Exhibit 1 shows this scientific research design for one of the media vehicles in the hypothetical test.

Third, we should expect to find small not large positive effect sizes of advertising on generating visits. Conversion studies support the following proposition: most persons responding with inquiries for free literature offered in advertising report having partial or complete plans to visit the destination. However, tourism advertising is more likely to generate other, very important, positive influences related to visitors receiving the free literature vs. not receiving the free literature.

The likely most substantial impact of tourism advertising is not on generating visits but in generating higher-quality visits than would otherwise happen. The findings of a quasi-experimental test support this proposition. Expenditures at the destination (Prince Edward Island, a Canadian Province), the number of activities, the perceived quality of the visit, the overall satisfaction with the trip, and the intention to return all increased substantially among visitors receiving vs. visitors not receiving the free advertising literature, a 130-page visitors’ information guide (see Woodside, MacDonald & Trappey, 1997). An unexpected additional finding in the PEI study was that the increases in visit quality influenced by the free literature were greater among repeat vs. first-time visitors. Possibly the free literature aided in getting repeat visitors to try new, unknown, or forgotten activities in the destination area in addition to their planned destination activities. Consequently,
the repeat visitors using the current free visitor’s guide had the highest dollar expenditures and greatest satisfaction with their visits compared to the other three groups (first time visitors using and not using the guide and repeat visitors not using the guide).

In addition, tourism advertising often has a small, positive, systematic impact on visits. Several rationales and findings from the advertising-to-sales literature support this proposition. Increases in advertising (from none to some advertising) often result in increasing unaided retrieval of the product advertised, and increases in unaided retrieval are associated with increases in sales (literature is reviewed in Woodside, 1996). Most linkage advertising (e.g. free visitor’s catalog) dramatically reduces the effort of competing and increases knowledge of how to complete many details in planning a visit to the destination. In fact, several studies in consumer psychology support the finding that the availability of such problem-solving tools influences choice decisions (see Bettman, Luce & Payne, 1998).

CONCLUSIONS

We use the term “golden rules” to generate discussion and cause action. A core conclusion from the meta-evaluation presented is that few explicit attempts have been made to explicitly improve the methods of evaluating of government tourism-marketing programs. The Australian 1993 audit deserves recognition as an attempt to create and apply a new paradigm for a tourism-marketing program evaluation. In part, our hope is that this study will serve as a foundation for additional attempts to transform the discipline.

Certainly, additional “golden rules” need to be developed. For example, scientific measurement of the long-run impacts of advertising on generating visits and attracting high-quality visits need to become standard practice in tourism-marketing program evaluations. Such scientific measurements of long-run impacts require creating additional treatment groups to enable measuring once only for each group – to avoid the introduction of bias from answering surveys.

Another golden rule would be to perform only comprehensive performance audits of government tourism-marketing programs. “Limited-focus” audits do not represent substantial savings in effort. The trade-off of accuracy and detail for greater speed results in unfounded and misleading confidence that a comprehensive audit is unnecessary. Combining, replicating, and extending the findings from qualitative and quantitative measurement is practical only in comprehensive audits. Useful comprehensive audits can be completed quickly (within two to three months), a time period not much greater than a limited focus audit.
REFERENCES


Assessing Alternative Methods of Performance Evaluation


Assessing Alternative Methods of Performance Evaluation


