

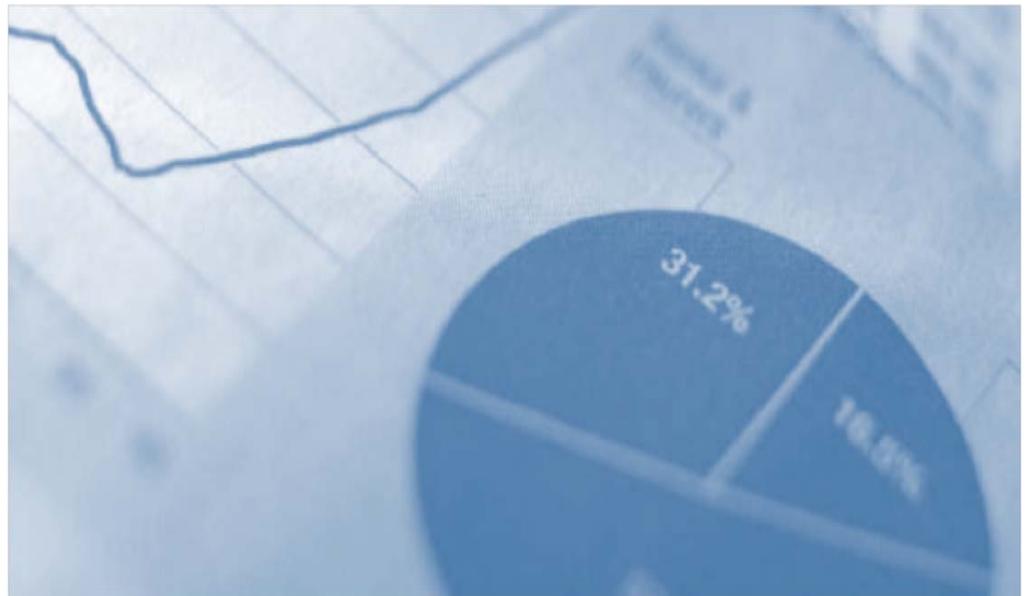
ROBERT WOOD JOHNSON FOUNDATION EVALUATION SERIES

An Ecological Understanding of Evaluation Use

A Case Study of the *Active for Life*
Evaluation

2010

By Judith Ottoson, Ed.D., M.P.H. and Diane Martinez, M.P.H.



ACKNOWLEDGMENTS

This report was prepared for the Robert Wood Johnson Foundation, under the direction of Laura C. Leviton, Ph.D., Special Adviser for Evaluation, by Judith M. Ottoson, Ed.D., M.P.H. and Diane Martinez, M.P.H. Judith M. Ottoson (jottoson@comcast.net) is an independent evaluation consultant and adjunct lecturer at San Francisco State University. Diane Martinez (dmartinez@rwjf.org) is a Rutgers Policy Analyst Fellow at the Robert Wood Johnson Foundation.

Copyright 2010 Robert Wood Johnson Foundation



Robert Wood Johnson Foundation

Route One and College Road East
P.O. Box 2316
Princeton, NJ 08543-2319

This publication is available for downloading from the Foundation's Web site at:
www.rwjf.org/pr/product.jsp?id=71148

Table of Contents

What is Evaluation Use?	3
Introduction	4
A Description of the Program Evaluated: <i>Active for Life</i>	5
A Description of the Evaluation	5
Case Study Methodology	6
Findings	7
- Types of Evaluation Use	9
- “Threads” of Evaluation Use	15
- Leveraged Knowledge	16
Facilitators of Evaluation Use	17
Barriers to Evaluation Use	18
Conclusions	19
Implications	20
<i>Appendix A: Literature Review on Evaluation Use</i>	22
<i>Appendix B: Full Description of Methodology</i>	25
<i>References</i>	33
<i>Bibliography</i>	35

What is Evaluation Use?

At its core, evaluations involve judgments about the *value* of something, such as a program or policy. Determining program value takes resources, time and effort. The result of this effort is usually a presentation, report or other products that feature findings about the program process or activities and events, and/or program outcomes that include the results or effects of program activities. Questions about evaluation use try to get at whether anyone ever did anything with those evaluation presentations, reports, or other products. Did anyone read them? Listen to them? Pass them on to others? Make a program or policy related decision? Change a program activity? Choose one course of action over another? Shift priorities? All of these might be considered kinds of evaluation use. To understand further how evaluators have explored evaluation use over time, see [Appendix A: Literature Review](#).

So, why do evaluators make such a fuss about evaluation use? Why do they keep studying it? Evaluation use is core to both evaluation standards¹ and practice². Deeper than these professional concerns, however, is core evaluation theory and belief that evaluations are supposed to alleviate the social problems to which programs and policies are directed (Shadish, Cook and Leviton, 1991). If no one ever does anything with evaluations or even knows of their existence, then resources, time and effort are wasted and opportunity for social betterment is lost (Ottoson and Hawe, 2009). Evaluation use is a social promise, not professional narcissism. Perhaps one of the respondents in this case study sums up evaluation use best: “The evaluation was as important as actually implementing the program.”

¹ [Program Evaluation Standards](#)

² [American Evaluation Association](#)

Introduction

Evaluations are supposed to be useful. As previously noted, both the standards of evaluation practice³ and evaluation theory⁴ support this assertion. For decades, studies have explored whether and how evaluations are used.⁵ Is there anything new to say now? We think so. We present here a case study of use of an evaluation of *Active for Life®: Increasing Physical Activity Levels in Adults Age 50 and Older*⁶ a program sponsored by the Robert Wood Johnson Foundation (RWJF). The breadth of both the program and the evaluation created additional ways of thinking about evaluation use—what it looks like, where you find it and how to connect it.

We found that the *Active for Life* evaluation was used extensively by multiple stakeholders. During the case study, we confirmed familiar kinds of use, such as conceptual or instrumental use (discussed below), as well as kinds of use not well described previously, such as *valuing use*.⁷ We also uncovered sequential patterns of evaluation use that we called “[threads](#)”⁸ and [leveraged use](#)⁹ of evaluation-related knowledge across time and contexts. These findings led us to the development of *The Ecological Model of Evaluation Use* ([Figure 1](#)); we use that model to tell the story of this case study.

This report seeks to be helpful not only to the Robert Wood Johnson Foundation and other foundations, but more broadly to the fields of evaluation, aging, physical activity and public health in understanding what evaluation use looks like in context and how it can be facilitated. The community use of evaluation may have particular interest for those engaged in community-based participatory research; the evidence-based program link to evaluation may have particular interest for policy-makers and researchers; the ecological understanding of evaluation use adds a complementary approach to familiar categories in the study of evaluation use.

“I do think there is a substantial amount of use. When looking at all the ways the evaluation data have been used and considered, it’s heartening and a bit surprising to see the range of ways in which Active for Life made an impact.”

Michelle Maloney, M.B.A.
Instructor Development
Director at Human Kinetics

³ The [Program Evaluation Standards](#) provide a guide for evaluating educational and training programs, projects, and materials in a variety of settings. The four program evaluation standards are utility, feasibility, propriety, and accuracy. Reference: [American Evaluation Association](#)

⁴ An ideal evaluation theory describes why selected evaluation practices lead to particular kinds of results across situations and contexts. Such a theory would include issues fundamental to program evaluation including social programming, knowledge construction, valuing, knowledge use, and evaluation practice. References: Shadish, Cook, & Leviton, 1991; [American Evaluation Association](#)

⁵ Studies of evaluation use seek to answer a basic question: Did anyone do anything with evaluation findings or process, such as make a decision or change a program? See [Appendix A: Literature Review](#)

⁶ [Active for Life](#)

⁷ Valuing use is the use of the core work of evaluation to place value on a program or policy; it is the use of the totality of the evaluation not solely its process and/or outcomes [Table 1](#).

⁸ [Threads of evaluation use](#)

⁹ [Leveraged use](#)

A Description of the Program Evaluated: *Active for Life*

From 2003 through 2006, the *Active for Life* program sought to implement two efficacious, evidence-based, physical activity interventions and to determine whether these interventions could be translated with appropriate adaptations to be effective in real-world, diverse, community settings. Nine lead organizations and 12 geographically diverse sites participated in this national program by implementing either Active Living Every Day (ALED) or Active Choices. (Publications of the evidence-based trials for Active Choices and Active Living Every Day are listed in the [Bibliography](#)). (Please click here for additional information: <http://www.rwjf.org/pr/product.jsp?id=37128>).

Active Living Every Day (ALED) was evaluated in two different formats, a 20-week program and a 12-week program. The original 20-week ALED program was conducted in a group setting with 12 to 20 participants aged 50 or older who were led by a trained facilitator. Active Choices was a six-month program that began with an initial face-to-face meeting followed by eight one-on-one telephone counseling sessions. Over four years, each intervention site aimed to recruit 900 adults, aged 50 and older who represented a variety of ethnicities and income levels. RWJF established the *Active for Life* national program office (NPO) at Texas A&M Health Science Center where NPO leadership provided technical assistance to the participating organizations and sites, while maintaining communication with RWJF staff, the program developers and the evaluators contracted by the Foundation. (Please click here for additional information: <http://www.rwjf.org/pr/product.jsp?id=37128>)

A Description of the Evaluation

RWJF commissioned the University of South Carolina Prevention Research Center to conduct an external process and outcome evaluation of *Active for Life* from 2003 to 2007.¹⁰ The process evaluation results revealed that the lead organizations could implement the two evidence-based interventions for older adults of diverse ethnicities and income-levels while still maintaining program fidelity with agreed-upon intervention adaptations. Open communication among the project staff, program developers and evaluation team was a key facilitator (Please click here for additional information: [“Results from the Active for Life Process Evaluation: Program Delivery Fidelity and Adaptations”](#)). The outcome evaluation findings revealed that similar to the previous efficacy studies, the program participants experienced statistically significant increases in moderate-to-vigorous physical activity and total physical activity, as well as decreases in depressive symptoms and stress, increases in satisfaction with body appearance and function and decreases in body mass index. (Publications with the results from the *Active for Life* program are listed in the [Bibliography](#)). (Please click here for additional information: <http://www.rwjf.org/pr/product.jsp?id=15205>).

¹⁰ On a parallel track to the process and outcome evaluations, the RE-AIM framework was introduced to help participating programs assess reach into the community and sustainability. (Please click here for more information on the RE-AIM framework: <http://www.rwjf.org/pr/product.jsp?id=15569>)

Case Study Methodology

The question for this case study focused on whether and how the *Active for Life* evaluation was used. To inform methodology, we conducted a review of the literature ([Appendix A](#)) to identify conceptual frameworks of evaluation use and its facilitation (Weiss, 1979; Leviton and Hughes, 1981; King and Pechman, 1984; Cousins and Leithwood, 1986; Alkin and Taut, 2003; Taut and Alkin, 2003; Johnson et al., 2009). Following this review, we developed an open-ended study protocol.

"I think your results do show use of the evaluation... of course I would love to see more examples, after all the time, blood, sweat and tears (yes, a few) that went into it. But you did conduct a number of interviews with key informants, so not sure doing more will yield anything more substantial."

Anonymous

"The conclusion seems to be that the AFL evaluation was useful because stakeholders made comments that indicated usefulness in each of the traditional categories"

Anonymous

Over a five-month period starting in 2009, we interviewed 23 informants identified through a snowball sampling technique. These hour-long recorded interviews explored potential evaluation users, evaluation products and communication strategies, examples of evaluation use or non-use, and facilitators of evaluation use. To capture diverse perspectives on the *Active for Life* evaluation, the interviewees included the RWJF staff, National Program Office (NPO)¹¹ staff, evaluators, grantees, program developers, consultants and other stakeholders. In addition, documents that provided evidence of evaluation use were collected, including e-mails, presentations, flyers and Web sites. Data analysis initially involved the application of literature-identified conceptual frameworks and evolved into *The Ecological Model of Evaluation Use*

([Figure 1](#)). A full description of the methodology can be found in [Appendix B](#).

Informant Feedback

The case study findings were reviewed and critiqued by the relevant RWJF personnel and sent to all 23 key informants. They were asked specifically about the plausibility of the use types and facilitators identified and the usefulness of the revised conceptual framework proposed. Such "member checking" adds to the veracity of case study findings (Merriam, 1988). Reviews of some key informants are included in this report.

¹¹ The RWJF National Program Office Model stresses the provision of technical assistance to its national programs as a way of ensuring that these grantees are well equipped to meet the challenges and opportunities that may be present in a grant initiative.

Findings

We introduce *The Ecological Model of Evaluation Use* to frame case study findings. Linear models just do not tell the interactive story we found. Like other ecological models, this one proposes multiple “eco-systems” or contexts of evaluation use in this case study, with multi-directional and multi-layered influences.

- The model centers itself on the core work of evaluation—*valuing*, e.g., determining if the program is “good” or “needs improvement.”
- *Valuing* embeds in the *immediate program*, *Active for Life* in this case, where multiple stakeholders judge the program. The immediate program includes both the program process, e.g., activities and events, as well as program outcomes, e.g., the results or effects of program activities.
- The *community* context encompasses the geographical settings of the program, e.g., Chicago or Baltimore; the organizations which house the programs, e.g., a local community-based organization or YMCA; past, current and future program participants; and local policy-makers, e.g., organizational administrators or local officials.
- The *field* encompasses professional and disciplinary interests and practices that directly or indirectly have a stake in the immediate program. In the case of *Active for Life*, the fields included physical activity and aging, and more broadly public health. Included in the fields are practitioners, academics, advocates and policy-makers. Private foundations straddle field and society influences.
- *Society* encompasses broad cultural, social, political, economic influences, trends and policies. Incorporated into these are social needs, values and priorities and the negotiations among them for public attention.

The multiple contexts, as represented by these concentric circles, influence each other. For example, through an evaluation the *field* might place *value* on the *immediate program* as a best practice and the publication of such findings might influence *society* (policy-makers) to provide additional funding for such *programs*.

“As Detroit was unique in many ways, I believe that your findings capture the grounded nature of our work with small, voluntary organizations to evaluate the AFL intervention. I can see our perspectives reflected in your aggregated findings. I commend your efforts, as each site worked with a unique set of circumstances and conditions and it can be a challenge, at times, to decipher the continuities. I feel that you captured the dynamics in our setting accurately and in an unbiased manner.”

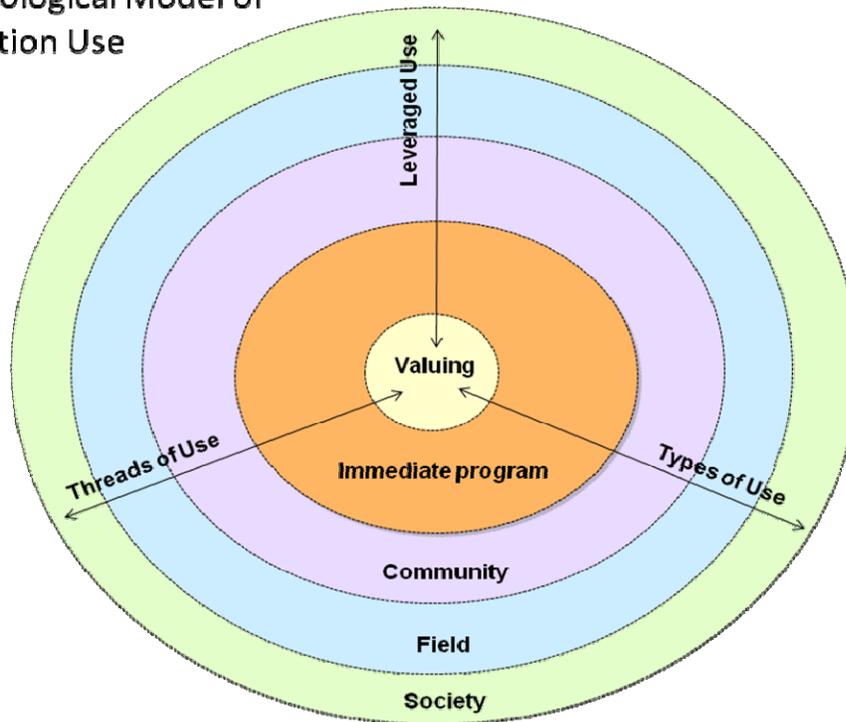
Kimberly Campbell-Voytal, Ph.D.
Program Director for Detroit ALED program, Assistant Professor at Wayne State University

Cutting across these concentric contexts of use are multiple paths to understanding evaluation use: types of use; sequences of use (we labeled “threads”); and leveraged use over time and context.

- We identified seven *types of evaluation use* in this case study including valuing use, instrumental use, conceptual use, evaluation learning use, symbolic use, communication use and decision-making use. These are discussed below and highlighted in [Table 1](#). As noted previously, some of these are familiar types of use, such as instrumental and conceptual; others are newly described, such as valuing use.
- The [threads of use](#) uncovered in the case study show sequences among types of evaluation use. Examples of such threads (page 15), are not meant to imply any universal order of evaluation use; rather, the threads suggest that types of evaluation use are not isolated events and understanding their connections may advance our understanding of how evaluations are used.
- [Leveraged use](#) includes the evaluation-related learning, tools and datasets, networks, relationships and other by-products of multiple evaluation stakeholders that influence not only the *immediate program*, but are carried across time, settings, fields, jobs and disciplines with influence elsewhere (described in detail on page 16).

Figure 1: The Ecological Model of Evaluation Use
(Source: Ottoson and Martinez, 2010)

The Ecological Model of Evaluation Use



Types of Evaluation Use

[Table 1](#) identifies, defines and exemplifies seven general types of evaluation use found in this case, as well as additional subcategories. While many of these types of use relate to the immediate *Active for Life* program, they are not limited to it and can be found in other contexts as suggested by the model and described below.

- **Valuing use.** Stakeholders used the evaluation to test the program, prove efficacy, provide evidence and offer credibility. Several informants indicated that without the ability to “test” the intervention they could not have changed, disseminated, marketed or recommended the intervention. While it may be tempting to dismiss this as symbolic use, e.g., the mere existence of evaluation, there is something different, deeper going on here. Valuing is a complex process that is as much about culture, social, personal, economic and political forces as it is about methodology. It is a process that is done in context amidst multiple stakeholders and power differentials. Fundamentally, **the evaluation enabled value to be placed on the whole** *Active for Life* program, as well as its components, its process and its purpose. Stakeholders did not use just the evaluation process or its findings; they used the *totality* of the evaluation effort to test the program.

“I liked this model – made sense to me and helped to integrate the complexity into a nice visual model.”

Sara Wilcox, Ph.D.
Evaluator, Professor at University
of South Carolina

“I think that would work nicely—it provides a framework...and everything is geared to the graphic/info as you have presented. But then...I’m a visual learner—so if someone questions that approach...I could be convinced otherwise. Think what you have shared gives clarity to a potentially confusing topic...so convoluted.”

Diane Dowdy, Ph.D.
NPO Deputy Director, Assistant
Professor at Texas A&M Health
Science Center

“Really liked it. Like how you explained fully but not too lengthy—providing links. Like how you were so open with process...trying to follow previous models but then having to take another twist or two to capture all the lessons. Easy to read...and I’m not an evaluator. Found the flow smooth—info helpful. Like how you emphasized the open communication and sharing...at all levels and from all parties.”

Diane Dowdy, Ph.D.
NPO Deputy Director, Assistant
Professor at Texas A&M Health
Science Center

Table 1: Types of Evaluation Use

TYPES OF USE	DEFINITIONS OF USE	EXAMPLES OF USE
VALUING USE n=13	Use of the core work of evaluation to place value on a program or policy; use of the totality of the evaluation not solely its process and/or outcomes	Test the program; provide evidence; prove efficacy; rigorous evaluation
INSTRUMENTAL USE	Evaluation process or outcome findings are directly tied to use, in relative proximity of time and with visible effect. Decision-making inherent in actions taken	
PROCESS FINDINGS		
Improve implementation n=6	Actions taken toward program activities and processes	Track activities; identify red flags; guide stakeholder implementation discussions; adjust program activities
Adapt program; program products n=7	Actions taken relative to program content, materials and core elements	Adapt, tweak, or modify program; identify essential program elements; add content; improve material usefulness
Improve evaluation n=1	Actions taken to modify the evaluation	Changes to evaluation process, instruments and fine tuning the system
OUTCOME FINDINGS		
Recruit n=6	Outcome findings used to meet program requirements for participant recruitment	Recruit participants and community partners; promote and market the program; use success stories and other data
Change program n=>10	Policy and structural changes to program (not tweaking)	Original 20-week program restructured into a 12-week option; medical release requirement dropped
Obtain Funding n=5	Outcome findings used to secure funding	Catalyst for funding; findings written into grant applications; publications attached

Note: In the first column the “n” equals the number of informants who provided an example of this type of use.

Types of use n= # of informants	Definitions of Use	Examples of Use
CONCEPTUAL USE	Evaluation leads to a change in understanding, not direct action	
Enable Choice n=5	Process and outcome findings enable choice among options	Choose processes; choose among program options; choose another program; make recommendations
Advance discourse n=4	Evaluation facilitates discussion and thinking, especially in the field	Provide insight; move the agenda, further the work; serve as a tipping point
Embed concepts n=4	Findings from evaluation embedded into other forms, context and fields	Lessons learned have other life in exercise guide; infuse finding in new program; supporting document; legacy effect
Provide reference point n=7	Findings confirm community tested, evidence-based program	Concrete example of a tested program, including its evaluation; source of data and methods; help build a “case” for the program
EVALUATION LEARNING USE	Changes in evaluation stakeholders and/or participating organizations as a result of learning that occurs during the evaluation	
Organizational learning n=6	Learning within the organizations that received program or evaluation funding; or that participated in the evaluation	The evaluation has taken the organization down a road of more in-depth evaluation; understand the importance of evaluation; sustained evaluation activities
Stakeholder learning n=4	Learning among diverse evaluation stakeholders including program managers, developers, participants and field associates	Insights; organic learning; understand implementation; how to get data back to scientific and participating communities
Evaluator learning n=3	Learning among participating evaluators	Manage large scale evaluation; more efficient; understand evaluation tool in context; realities of data collection

Types of use n= # of informants	Definitions of Use	Examples of Use
SYMBOLIC USE n=2	The mere existence of an evaluation (not its results) is used to persuade or to convince	Evaluation existence aided participant recruitment; organizational and/or personal credibility for evaluation participation
COMMUNICATION USE	Evaluation findings are used as they move through multiple communication channels	
Advocacy n=7	Evaluation findings used to influence decision-makers and resource allocation	Educate lawmakers; testify to own physician; attend Congressional briefing on findings
Professional dissemination n=7	Findings used in multiple professional dissemination channels	Publish findings in scientific journals; conference presentations; use data in dissertation; written reports
Networks n=3	Interagency networks around program area, such as aging or physical activity	Contribute ideas to evaluation instrument; share learning from evaluation process, not just outcomes
Community events n=5	Evaluation finding used in community events	Evaluator shares findings with participants and community; participant success stories; sense of pride; reinforce new habits; accountability moment; recruitment
Web sites n=4	Electronic access to evaluation findings, tools and processes	Multiple stakeholder Web sites; Web site hits can lead to personal contact for more findings
Media, press n=1	Evaluation findings used in media releases	Press release to newspaper; findings published in 5 county magazine
DECISION-MAKING USE	Use of the evaluation to inform decisions	Improve program process; recruit participants; structural change to program; advocate; disseminate

- **Instrumental use.** Stakeholders directly used process and outcome evaluation findings to take concrete, near-term action. Process findings were used to improve program implementation, adapt the program or its products, e.g., change manuals and to improve the evaluation, e.g., refine tools. Stakeholders used outcome findings to recruit future program participants, change the program structure, e.g., reduce program duration and to obtain funding for other programs. Some of these actions occurred in the ecological context of the *immediate program*, e.g., implementation changes; some changes occurred in the *community*, e.g., participant recruitment; some changes occurred in the *field* of physical activity, e.g., securing funding for other projects.
- **Conceptual use.** The *Active for Life* evaluation informed some stakeholder understanding, even if they took no direct action. The evaluation enabled choice among alternatives, e.g., program options or processes; it advanced discourse in the field by moving the agenda forward or providing insight; it embedded findings in other forms or fields, e.g., infused into a new program or exercise guides; it provided a reference point for action or decisions, e.g., case building for other programs. In the ecological model, we found many of these conceptual examples in the “field,” particularly in physical activity and aging.
- **Evaluation learning use¹².** Various stakeholders and organizations learned about evaluation by participating in the totality of the *Active for Life* evaluation effort. Some organizations that participated in the *Active for Life* program sustained evaluation activities after funding subsided or initiated evaluations of other programs. Diverse stakeholders, including program managers, developers and participants, used the evaluation experience to gain insights or understand processes differently. Even the evaluators learned more about the management of large-scale databases or the realities of data collection. Examples of this type of use emerge in the ecological model in the contexts of the immediate program, community and field.
- **Symbolic use.** The conduct of an evaluation to make a program, funder or stakeholder *look* responsive, rather than *be* responsive to program results, represents symbolic evaluation use. A few examples of this kind of use emerged in the case study. In these examples, respondents referenced or touted the *Active for Life* evaluation to gain personal, programmatic or

“I needed more clarity about how one gets to a strong conclusion of usefulness. Is it the breadth of comments? The nature of the comments? An interpretation that comments are evidence of impact?”

Anonymous

“I may be biased...but think that you clearly showed extensive use of evaluation...on multiple levels and in multiple ways. I was surprised as to how you captured so many ideas and expressed them so succinctly.”

Diane Dowdy, Ph.D.

NPO Deputy Director, Assistant Professor at Texas A&M Health Science Center

¹² Rather than using the term “evaluation process” as a category of use (Patton, 1997), we have labeled a comparable kind of use as “evaluation learning.” The re-labeling helps avoid confusion between “using process evaluation findings” and “learning from the totality of the evaluation effort” (not just its process).

organizational credibility by virtue of being associated with an evaluation. Symbolic use in this case was most often tied to other kinds of evaluation use and not a sole example of use.

- **Communication use.** Previous studies of evaluation use established the education of decision-makers as a category of evaluation use (Cousins and Leithwood, 1986). For this case study that category morphed into the broader category of communication use to capture advocacy, professional dissemination, networking, community events, Web sites, media, resource centers, intra-organization reporting and education. These were not sterile communication channels through which evaluation learning and findings passed untouched. These were implementation channels through which evaluation learning and findings were adapted, adopted, engaged, discussed and *used* (Pressman and Wildavsky, 1984). For example, some communities used evaluation findings to testify locally and to celebrate their collective accomplishment; other stakeholders took the evaluation findings to testify on Capitol Hill. In these communications, stakeholders engaged and used the evaluation.
- **Decision-making use.** Stakeholders used evaluation for decision-making in nearly all the preceding types of evaluation use. Stakeholders made decisions to evaluate (valuing use or symbolic use); decisions to improve the program or its implementation (instrumental process use); decisions to recruit or change the program (instrumental outcome use); decisions to choose or embed concepts (conceptual use); and decisions to communicate about the evaluation (communication use). With decision-making inherent in all of these categories, we were loath to pull it out as a separate category or put it on par with actions taken. In the end we decided to do so, however, so that the use of evaluation for decision-making not be embedded out of sight and out of mind in this case. Stakeholders used the evaluation for decisions and action.

“Yes—makes good sense and I think helps understand the multi-level approach and the openness to using the lessons learned. Isn’t that exactly what you are trying to demonstrate with this piece? --don’t be tied to following a format...but have an open mind to learning what new there is to learn while being thankful for validating previous lessons learned? I liked it...and tickled that AFL could demonstrate that so clearly (especially that you picked up on the uniqueness of learning).”

Diane Dowdy, Ph.D.
NPO Deputy Director, Assistant
Professor at Texas A&M Health
Science Center

These individual categories help illuminate ways in which various stakeholders used the *Active for Life* evaluation and the multiple contexts in which they did so. Although valuing is core to the work of evaluation, the category of “valuing use” offers a new contribution through this case study, as does the expansion of “process learning” to “evaluation learning” and “education use” to “communication use.” As helpful as these categories are, they told only part of the story on evaluation use for this case study.

“Threads” of Evaluation Use

To more fully tell the story of use we found that we needed to move away from a sole focus on separate categories of use. Our struggle with what to do about decision-making use linked to action or other kinds of use led us to explore patterns or links among categories. We labeled these as “threads of evaluation use.” These threads involve tracing what precedes, intervenes with and/or follows types of evaluation use across contexts, time and stakeholders, as shown in the ecological model. Some examples of evaluation use threads found in this case study are illustrated below in quotes from informants where the individual categories of use have been italicized:

- If programs had just been implemented, but not evaluated (*valuing use*), they wouldn’t have been picked up (*decision-making*) since the push now is to get on the evidence-based list (*reference point*) by having published peer-reviewed articles (*professional dissemination*)... being on the CMS [Centers for Medicare and Medicaid Services] listing of evidence-based programs (*reference points*) made the program eligible for reimbursement (*obtain funding*).
- We took the program and grew it (*change program*) in so many different directions (*adapt program*) because we knew it worked and could prove it worked (*valuing use*).
- Having these articles (*professional dissemination*) gave us leverage to present at conferences (*professional dissemination*) since the findings carry a lot of weight (*valuing use*) ...can reach a lot of thought leaders (*decision-making*) and people that can implement the programs at the community level (*improve implementation*).
- We had the luxury to experiment (*evaluation learning*) with a shorter course (*change program*) and test (*valuing use*) whether the shorter program was as effective.

Valuing use anchors each of these threads. Through evaluation learning and dissemination, decisions are made, program changes occur, and funding may be attracted. These kinds of use crisscross multiple contexts of evaluation use and are characteristic of the reciprocal determinism or mutual influence inherent in ecological models (Green, Potvin and Richard, 1996). For example, the field can influence the immediate program by recommending promising practices and the immediate program can influence the field by demonstrating the feasibility of practices. By understanding the evaluation use categories, it is possible to push a further understanding of how the categories relate to each other, over time and through multiple contexts. In the end we did not find one story of evaluation use, we found a web of interconnected vignettes.

Leveraged Knowledge

Much of the evaluation use literature focuses on use in the context of a specific program, its organization, or its stakeholders. The wide canvas of the *Active for Life* case study allowed—demanded—an exploration of use of evaluation learning beyond the immediate program. Stakeholders changed jobs, evaluation findings diffused, evaluation tools were adapted to other programs and insights kept emerging. In all of these, the evaluation experience from the *Active for Life* program was leveraged over time, context, programs and stakeholders. Below are some examples of “other” situations in which the *Active for Life* evaluation was used:

“Have the results of the evaluation been used in other areas in addition to physical activity and aging? . . . If there are examples that could be cited, I think it would make for a stronger case for evidence of evaluation use. [See RE-AIM Example]”

Teresa Keenan, Ph.D.
Health Research Team Leader
at AARP

Other programs:	Evaluation tools, forms and database served as a model for evaluation of a new program; improve evaluation of other programs
Other grants:	Evaluator used the <i>Active for Life</i> evaluation process to identify core program components in a new NIH grant
Other sites:	Evaluation process and tools used at expanded ALED program sites; evaluation process infused into other programs; evaluation tools used in other states
Other purposes:	Evaluation learning led to RE-AIM expansion from summative to formative evaluation tool
Other fields:	Evaluation learning led RE-AIM to move beyond behavioral medicine and public health to include the fields of aging and physical activity
Other organizations:	Networks of professionals or organizations leveraged evaluation learning to multiple participating organizations
Other clients:	Program developers sold improved participant manuals to other clients working in related fields
Other participants:	Findings of effectiveness opened the door to working with more diverse communities
Other work:	Reshuffling of stakeholders over time led to new use for findings, e.g., local program managers work together on new grants; program manager serves as an evaluation consultant

To focus evaluation use only on the immediate program or context misses a lot. While the threads of evaluation use tell a story, an exploration of the leveraged evaluation-related knowledge paints the big picture.

Facilitators of Evaluation Use

The Johnson et al. (2009) framework informed, not dictated, our findings of facilitators of evaluation use.

“Did find it interesting that ongoing use of data was helpful...but then a double-edged sword in that it took a while to get the cleaned data/analyzed/publication before could be fully utilized by the sites.”

Diane Dowdy, Ph.D.
NPO Deputy Director, Assistant
Professor at Texas A&M Health
Science Center

- ***Open communication and stakeholder involvement.*** After interviewing 23 stakeholders involved in the *Active for Life* program, it was quite telling that everyone, e.g., NPO staff, evaluators, lead organization staff, RWJF staff, was in agreement as to the purpose of the *Active for Life* program and its evaluation. Open communication and stakeholder involvement established a clear understanding of the objectives of program goals, program processes and measurable evaluation outcomes. The grantees felt comfortable in helping the researchers understand how the *Active for Life* programs could be translated into real-world settings and what adaptations needed to be made to enhance participant recruitment and commitment to the entire length of the Active Living Every Day (ALED) and Active Choices programs.
- ***Competence of the evaluators and the evaluation quality.*** Multiple informants expressed their appreciation that the evaluation team at the University of South Carolina conducted the *Active for Life* evaluation with a flexible approach to learn how best to collect the evaluation data from the lead organizations. One informant commented that “the evaluator knew and understood how to work with community-based organizations.” Mutual respect between the evaluators and program staff enabled the ***grantees to become more receptive to the evaluation process.*** This evaluation format ultimately enhanced the completeness of the evaluation data and ***broadened the grantees’ understanding of how the evaluation findings could be used*** to improve recruitment efforts and to create new partnerships for additional locations offering the ALED or Active Choices program during the four-year grant period.
- ***Preliminary evaluation findings shared on an ongoing basis during the annual grantee meetings.*** This strategy for disseminating evaluation findings further facilitated use and understanding of the evaluation results by the lead organizations. One informant commented that preliminary evaluation results during the grant period were highly anticipated because the lead organization staff and program participants wanted to know if the evidence-based program worked for them. Buy-in at the local level was achieved at one site particularly well since they had the following approach to the *Active for Life* program and evaluation: “Let’s participate so that we can show the researchers [of the original efficacy studies] how it works and doesn’t work for us.”

- **Article publications and conference presentations.** Use of the evaluation findings by professionals in the fields of aging and physical activity resulted from article publications and conference presentations given by the *Active for Life* NPO staff and evaluators. One informant from the field viewed the *Active for Life* evaluation results as being a “tipping point” for certain national organizations to provide more financial support for evidence-based programming at the local level. The *Active for Life* evaluation was also viewed as “truly a landmark in the level of evaluation” whose full benefit was beyond just scientific papers. Another informant commented that: “From the national perspective, if [the *Active for Life* grantees were] successful in implementing the programs and if the evaluation data was strong, it would help [certain national organizations to] become more engaged in evidence-based programs. The evaluation was as important as actually implementing the program.”

Barriers to Evaluation Use

While this case study has presented a range of use examples for the *Active for Life* evaluation, the informants also reported barriers to use.

- **Time.** One of the biggest impediments to the use of evaluation findings was the time it took for the grantees to receive them during the four-year grant period. Certain stakeholders viewed the evaluation data as being crucial for their sustainability efforts in securing future funding, with the delays preventing them from having data to support the observed successes that were occurring on the ground. Since the evaluators published the final evaluation results in 2008, two years after the grant had formally ended, some interviewed stakeholders were not aware of their existence in a peer-reviewed journal.
- **Non-use of the evaluation.** Although one Active Choices grantee, Blue Shield of California, had a successful experience in implementing the program, a corporate decision was made to eliminate the entire department that ran the *Active for Life* program. Despite the positive evaluation results, this was an example of how a corporate decision, not the evaluation findings, determined the program’s future by ending it.
- **The Foundation’s shift away from the field of aging.** In 2005, the Foundation moved towards its new commitment to reversing the childhood obesity epidemic. This shift occurred about half-way through the *Active for Life* program and resulted in some stakeholders feeling like “RWJF had abandoned the aging field” once the *Active for Life* program ended in 2006. The Foundation’s shift away from aging may have prevented the *Active for Life* evaluations from being used if potential users may not have viewed it as a go-to-source for this information.
- **Partnerships may have ended after the grant period; sustainability was not possible with lack of resources.** One informant commented that when leaders at a university and/or community-based organization leave, the memory and institutional knowledge of the program’s success fades for that organization. As a result, the momentum to sustain a program may end and the previous partners and access to resources may no longer be available.

Conclusions

At the beginning of this report, we stated our conclusion that the *Active for Life* evaluation was (and is still being) *used*. By this point, the reader might agree with us. The number of examples, their specificity, the corroborations and documents that support them, and the number of stakeholders who offered them convinced us that *use* happened. Further, we conclude that there was a lot of use—multiple examples, by multiple stakeholders, in multiple contexts—and that some of this was “big” use, e.g., changed program structure, organizational learning, advocacy and leveraged use to other programs. Understanding and discovering categories of evaluation use enabled us to hear the stories or threads of evaluation use and see the bigger picture of leveraged use. Finding the obvious, *valuing use*, reminds us that evaluation use occurs not just in categories, but also in the totality of the evaluation. This core work of evaluation—valuing—anchors all use.

The facilitators and barriers to evaluation use helped explain why use occurred in this case. Further, they concur with the evaluation use literature. Johnson et al. (2009) concluded that: “...engagement, interaction and communication between evaluation clients and evaluators are key to maximizing the use of the evaluation in the long run (p.389).” These key actions existed and persisted in this case study.

Implications

- **For the Robert Wood Johnson Foundation.** The *Active for Life* program is an exemplary model for how the communication quality and partnerships among RWJF staff, NPO staff, project staff, and the evaluators were essential to use of the evaluation findings. Certain informants commented that by having RWJF staff and NPO staff set the tone for the complementary roles of program implementation and evaluation at the beginning of the initiative, the program staff and program participants better appreciated the evaluation process and learned how the results could strengthen their programming and potential sustainability. RWJF program staff should look to the *Active for Life* evaluation as exemplifying the importance of open communication, strong dual leadership from the NPO and evaluators and the importance of engaging stakeholders at the start of program evaluation. (Please click here for another RWJF-funded evaluation product: "[A Guide for Engaging Stakeholders in Developing Evaluation Questions](#)")
- **For foundations in general.** Foundation staff, program staff and evaluators must be in constant communication and support each other in holding both program implementation and evaluation as being complementary and important for learning and program improvement. Mutual respect among these groups will formulate strong leadership from the beginning that will carry through the grant period and enhance data collection and use/understanding of the evaluation findings. Since foundations do not always financially support professional networks, it was recommended by some informants that foundations recognize the benefits of professional networks and consider supporting these networks to sustain and further advance a particular field. It was also recommended that foundations encourage the publication of evaluation findings beyond peer-reviewed journals, e.g., trade magazines.
- **For evaluators.** An ecological understanding of evaluation use opened the door to multiple paths to study evaluation use, including types of use, threads of use and leveraged use. It gave us a broad perspective. For the visual types—whether evaluators, policy-makers, program managers, or community members—the model provides a way to see, track, connect, place, ground and envision evaluation use. We think the model is worth testing further. This brings us to the naming of *valuing use* as another kind of use to add to categories of evaluation use. In this case, evaluation was used not just for its components, e.g., process learning or outcomes, nor for its symbolic presence. The totality of the evaluation tested and valued the program. *Valuing use* anchors all evaluation use. In addition, evaluator competence and understanding of the evaluation users are important during the evaluation process to ensure that findings are relevant and understood by the users. The *Active for Life* evaluation was effectively used to sustain some of the programs since there was open, honest communication among grantees, the NPO, the program developers and the evaluators on what worked, and what didn't work, and how this could be remedied.

- **For other stakeholders.** Evaluation findings can be useful for the program developers and private marketers of the program in order for them to know how the program is translated into real-world settings; and for marketers, how to better communicate about the program. Not every program will work in every setting. By testing and understanding adaptations that can be made while maintaining program fidelity, the developers learn what works and the marketers learn how to better serve their customer base before, during and after the programs have been implemented. Professional conferences and networks also serve as a means for bringing a new and evolving life to evaluation findings and their use.

In closing, this case study of use of an evaluation of the *Active for Life* program provided new insights, some new categories of evaluation use and an ecological approach to the study of evaluation use.

The case study also confirmed some of the recent work in the evaluation literature, especially on the facilitators to evaluation use. We hope we accurately reflected the experience of case study participants. Mostly, we hope this case will be useful beyond these pages.

“Struck me as a substantial amount of use (which was so nice to see!!)”

Sara Wilcox, Ph.D.
Evaluator, Professor at University of South Carolina

“It was not a surprising amount... given we were really “living” that experience. But, it was great to read the report and be reminded of just how much was done and how much the evaluation was used (which makes me feel great about the work).”

Sarah Griffin, Ph.D.
Evaluator, Assistant Professor at Clemson University

Appendix A Literature Review on Evaluation Use

Before assessing whether or not *use* of the *Active for Life* evaluation occurred, an understanding of the theory, practice and methodology of evaluation use was sought by conducting a literature review of empirical research published during the last 40 years. The following online research databases were used to identify articles for this literature review: JSTOR, PROQUEST and ERIC. The following key words were searched for in the databases: *evaluation use*, *evaluation utilization*, *use of evaluation* and *process evaluation*. This literature review resulted in an examination of 36 articles published from 1979 to 2009. A summarization of the evaluation use concepts outlined below enabled us to take what we learned from the literature and apply it to our evaluation use case study methodology.

Types of Evaluation Use

The concept of “use” is core to both the foundations of evaluation theory (Shadish, Cook and Leviton, 1991) and program evaluation standards (The Joint Committee on Standards for Educational Evaluation, 1994). Evaluations are supposed to be useful. How evaluators have understood use and what facilitates it have evolved over the last few decades. One of the early understandings of use identified seven different categories of evaluation use, ranging from the rationale of problem solving to the fodder for political negotiations (Weiss, 1979). By the early 1980s, three general kinds of use of evaluation findings were identified: instrumental, conceptual and symbolic (Weiss, 1980; Leviton and Hughes, 1981; King and Pechman, 1984). Instrumental use occurs when evaluation findings have been used in decision making or problem solving. Such use is more likely to occur sooner rather than later and directly rather than indirectly. Conceptual use occurs when a person has been enlightened by the evaluation findings, but no direct action is taken. Symbolic use occurs when the evaluation findings are not used, but rather the existence of the evaluation is referenced to persuade or convince. In 1997, Patton introduced the concept of *process use* in which individuals alter their behaviors and/or organizations change their program structure as a result of learning from the evaluation process. This added lens broadened an understanding that an evaluation can be useful for more than its findings.

In subsequent literature reviews, new and refined understandings of evaluation use were identified. Cousins and Leithwood (1986) examined 65 empirical studies published between 1971 and 1986. These researchers highlighted the following three concepts of use of evaluation findings: (1) use as decision making, (2) use as education, and (3) use as the processing of evaluation information. Examples of decision-making included discrete decisions about program funding, program operations and/or program management. With respect to the concept of use as education, evaluation findings may help educate decision-makers and program staff about the structure and intended outcomes of the program being evaluated. The processing of evaluation information occurs when program staff responds to the evaluation findings by either acting upon the recommendations and/or improving aspects of the program. Another literature review found that knowledge-for-action theories—transfer, translation, diffusion, implementation and knowledge utilization—shape what is valued, how it is valued and what constitutes evaluation use

(Ottoson and Hawe, 2009). A further expansion beyond direct use of evaluation processes or findings was conducted by Kirkhart who established the theory of *evaluation influence*, which she defined as the capacity of persons or things to shape others in indirect ways (2000). Researchers Henry and Mark (2003) have also explored this concept, viewing the differing paths and levels of *evaluation influence* as a means through which evaluations lead to the improvement of social conditions.

Facilitators and Barriers to Evaluation Use

The concept of use shapes the understood (and researched) barriers and facilitators of such use. In 1986, Cousins and Leithwood identified the following six factors that affected evaluation implementation in the empirical studies they reviewed: evaluation quality, credibility, relevance, communication quality, findings and timeliness. These researchers also highlighted the following six factors related to decision or policy settings: information needs, decision characteristics, political climate, competing information, personal characteristics and commitment and/or receptiveness to evaluation. After reviewing the evaluation use literature, Cousins and Leithwood concluded that evaluation use was most evident when intended users of the evaluation were involved from the beginning, the evaluation methods were appropriate and the evaluation findings were relevant to the intended users. Non-use and diminished use were also potential possibilities if the evaluation results were perceived by the intended users as a threat (Cousins and Leithwood, 1986).

Researchers Alkin and Taut (2003) and Taut and Alkin (2003) further contributed to the literature through their exploration of the following barriers to evaluation use: human factors (evaluator and user characteristics); program context (contractual obligations, fiscal constraints, organizational characteristics, external community factors); and evaluation characteristics (methodology, communication of findings, personal characteristics, and actions of the evaluator).

In a comprehensive review to identify facilitators of evaluation use, Johnson and her colleagues (2009) examined 41 studies of evaluation use from 1986 to 2005, and classified them into the following three categories: (1) evaluation implementation, (2) decision or policy setting, and (3) stakeholder involvement. [Table 2](#) illustrates 22 specific characteristics of these three overarching categories.

After examining the literature, Johnson and colleagues found that the most commonly referenced characteristics of evaluation use were the following: “stakeholder involvement with commitment or receptiveness to evaluation;” “communication quality;” and “personal characteristics” of users. In addition to emphasizing the importance of engaging stakeholders in program evaluations, the researchers also highlighted the importance of evaluator competence in facilitating use of evaluations (Johnson et al., 2009).

Table 2: Facilitators of Evaluation Use

Evaluation Implementation	Decision or policy setting	Stakeholder Involvement
Communication quality	Personal characteristics	Involvement with commitment or receptiveness to evaluation
Timeliness	Commitment and/or receptiveness to evaluation	Involvement with communication quality
Evaluator competence	Political climate	Direct stakeholder involvement
Evaluation quality	Decision characteristics	Involvement with credibility
Findings	Competing information	Involvement with findings
Relevance	Information needs or the evaluation audiences	Involvement with relevance
Credibility		Involvement with personal characteristics
		Involvement with decision characteristics
		Involvement of stakeholders facilitated the introduction of their information needs

(Source: Johnson et al., 2009)

Appendix B Full Description of Methodology

Active for Life was a large-scale evaluation covering multiple sites, thousands of program participants and multiple layers of stakeholders. This broad evaluation canvas needed a methodology that would allow a holistic exploration of evaluation use. Case study methodology is selected as “...the preferred strategy when ‘how’ or ‘why’ questions are being posed, when the investigator has little control over events and when the focus is on a contemporary phenomenon within some real-life context” (Yin, 2003a, p1). Our guiding research questions asked “how” evaluation was used (or not) and “why” use was facilitated or hindered. Case study investigators had no control over events. Begun in 2003 and ending in 2007, the evaluation was sufficiently contemporary since stakeholders, evaluation products, stakeholders and links were possible to find and remember. Evaluation of *Active for Life* programs at some sites and evaluation publications continue today despite completion of Foundation funding.

Having settled on the case study methodology, it was necessary to pick among several approaches that best fit this study (Merriam, 1988; Stake, 1995; Yin, 2003a; Yin, 2003b). These approaches vary by the types of fields in which they were developed (e.g., education or policy), theoretical approach (e.g., developing theory from the ground up versus starting with a conceptual framework) and methods proposed (e.g., only qualitative versus a mixed method approach). The approach selected as most appropriate for this case study was that of Yin (2003a) because of its multi-disciplinary applications, the match to the questions asked and context given and investigator familiarity with this approach (Ottoson, Rivera, DeGroff, Hackley and Clark, 2007).

Yin (2003a) identifies five components or levels of case study research design: (1) study questions; (2) study propositions (conceptual framework); (3) unit of analysis; (4) logic linking data to the propositions (data collection); and (5) criteria for interpreting findings (analysis). Each is reviewed below for their application to this case study.

Study Questions

The guiding study questions for this case study were as follows:

- Was the *Active for Life* evaluation—not just its findings—useful?
 - If yes, what were the parameters of use, e.g., for whom? when? how?
 - If not, why not?
 - What constituted evaluation “use” in context?
- How can findings from this and other evaluations sponsored by the Robert Wood Johnson Foundation be made more useful?
- What are the general implications of this evaluation use case study on the *Active for Life* program for foundations, leaders in the fields of aging, medicine and public health, as well as other stakeholders?

These questions were used as the entrance to the literature review, the basis for the study protocol, and the organization for the case analysis and reporting. The questions were approved and supported in their exploration by the Foundation.

Conceptual Framework

Having completed a literature review on evaluation use (Blake and Ottoson, 2009), as well as written and theorized about evaluation use over time (Ottoson and Hawe, 2009), we agreed with Yin about the importance of making the investigator’s conceptual approach transparent. Recent literature reviews identified several frameworks that have been employed to classify and understand types of evaluation use, as well as facilitators and barriers to use. We initially settled in this study on three of those frameworks previously described in the literature review. First, was the classification of evaluation use as instrumental, conceptual, or symbolic (Weiss, 1980; King and Pechman, 1984; Leviton and Hughes, 1981). Second, was the classification of evaluation process use (Patton, 1997). Third, was the classification of evaluation use as decision-making, education and the processing of evaluation information (Cousins and Leithwood, 1986). Working within the consistency of each framework and recognizing the potential redundancy across frameworks, we found no compelling reason to choose one framework over another in the initial exploration of evaluation use. We wanted as many lenses as were appropriate to the questions¹³.

In exploring the facilitators and barriers to evaluation use, we settled on the framework developed by Johnson et al., 2009. As noted in the literature review and in [Table 2](#), this framework identifies three broad areas of facilitators, i.e., evaluation implementation, decision or policy setting and

¹³ We agreed with Johnson, et al, (2009, p379) on the impracticality of adding “influence” to this case and stayed with the concept of use.

stakeholder involvement. A total of 22 subcategories of facilitators fall within these three broad categories.

While we agreed with Yin's call for conceptual transparency, we reserved the analytical and experiential opportunity to challenge these frameworks. If the data did not fit a category, suggested another one, or took us in some other direction, we followed it.

Unit of Analysis and Case Study Sample

The unit of analysis or "case" is use of the process and outcome evaluations of the *Active for Life* program funded by the Robert Wood Johnson Foundation. There were several other evaluations occurring in the same context and around the same time as the case-focused evaluation. These included an AARP marketing and policy initiative; a comparison of the Active Choices and Commit to Be Fit programs; the program developer's evaluation of program materials; and the introduction of RE-AIM¹⁴ into the *Active for Life* program. While it was possible to separate the first three evaluation efforts from the case study, it was not practical to do the same with RE-AIM. Although this framework was a part of the *Active for Life* evaluation from the beginning as a summative tool, half-way through the evaluation it began to be used as a formative tool that aided planning and evaluation.

Within the case, we sought a snowball sample of potential users, products and transmissions of evaluation processes and findings. A snowball sample is a type of purposive sampling in which early key informants identify others who meet the same criteria for inclusion, hence increasing the size of the "snowball" (Rich, 1977; Leviton and Boruch, 1983; Merriam, 2009). Starting with the Foundation's evaluation leadership and program management, the *Active for Life* National Program Office (NPO), and the process and outcome evaluators, the sample reached out to participating organizations, leadership in the fields of aging and physical activity, government agencies, frontline program managers, communication experts and program developers. We were referred to program participants, but unsuccessful in reaching them. A total of 23 stakeholders were included in the final sample.

One note of explanation needs to be made about this sample. Because of the passage of time, many of those interviewed were reshuffled into different roles, contexts and programs by the time the interviews were conducted. For example, someone formerly in Foundation management is currently in the federal government and another Foundation employee is now in private consulting; a local program director remains in the same organization, but moved to another state; and program management has played subsequent consulting roles in the aging field. As a result, a number of stakeholders brought more than one perspective to bear on use of the *Active for Life* evaluation.

¹⁴ RE-AIM is not a theory or conceptual model, rather it is a framework and a set of criteria for evaluating interventions that are intended to eventually be broadly implemented or widely adopted. The acronym stands for reach, efficacy, adoption, implementation, and maintenance. (Please click here for additional information: http://re-aim.org/2003/FAQs_support.html)

Data Collection

Data were primarily collected through 20 telephone interviews using an open-ended protocol. (Two interviews had more than one respondent on the call). The interviews were conducted between December 2009 and April 2010, each lasting approximately one hour. Both investigators participated in all interviews.

The protocol used to guide the interviews was based on an extensive review of *Active for Life* program materials and publications, as well as the literature review. The protocol, on the last page of Appendix B, explored issues around evaluation use in the case study, as well as identification of relevant evaluation products (e.g., reports or flyers), evaluation transmission strategies, (e.g., Web sites or conferences) and other informants we might include in the snowball sample. While the protocol was originally developed for program leadership, it turned out to be useful and appropriate for all types of stakeholders, with only minor changes. The protocol was sent to informants before the scheduled interview so that they were familiar with the questions. Every informant did not answer every question.

Two points need to be made about the protocol. One is that we did not define “evaluation use” to key informants and the other is that we did not prompt or solicit answers about any particular type of evaluation use. Since the literature does not offer a single understanding of evaluation use, we decided not to impose such a limitation on informants. They were able to identify and explain evaluation use as they understood it to fit within their context and this case study.

In the course of interviews, additional and related questions were generated. In some cases these were “call forward” questions to be asked of future informants. In some cases these were “call back” questions to be clarified with those already interviewed. Most often these questions were around confirming an understanding of evaluation use.

All calls were voice-recorded with permission of the informant. Both of the investigators took extensive notes, with intent of capturing respondents in their own words as much as possible. Without resources to transcribe the interviews, we used the extensive dual notes taken as the data source, with the voice recordings as clarification for any uncertainty or discrepancy between notes. One investigator typed her notes in-depth shortly after each call, in some cases transcribing extended answers to questions. The other investigator’s notes and the voice recordings were used as a check on the more complete set of notes.

In addition to interview notes, we collected artifacts or documents where possible to confirm examples of evaluation use. These included publications, e-mails about meetings and exchanges, and recruitment flyers. Some respondents sent follow-up e-mails to the investigators to add other understandings of evaluation use, directed us to Web sites, or added further clarification to their interview comments.

Data Analysis

Finding types of evaluation use was the first data analysis priority. While the conceptual frameworks alerted us to kinds of evaluation use and informants provided their self-identified examples of use, the following question was continuously explored during the data analysis process: *What counts as evaluation use in this case study?* First, any category or type of evaluation use identified needed to be corroborated by more than one key informant to be included in the case. (This was true for all categories except two which could be corroborated by other means.) Second, the informant needed to provide a description of their example of evaluation use in context. Descriptions that were about the *Active for Life* program, not the evaluation, were eliminated from consideration. Third, a use example was considered stronger if it was identified by stakeholders in more than one role, such as an evaluator and a program manager. Fourth, use examples were considered stronger if they were supported by artifacts, such as e-mail, Power Point presentations, or flyers. Lastly, if informants identified another key informant to back an assertion of evaluation use and the second informant failed to do so, the example was dropped.

Analysis began by reviewing interview notes across types of key informants, such as program managers, developers or evaluators. The intent was to allow the understanding of evaluation use to be shaped simultaneously by all stakeholders, not sequentially by one type then another. The in-depth interview notes were read first and potential examples of evaluation use, facilitators and barriers to use, and recommendations about use were identified and put into separate Excel spreadsheets. One investigator worked primarily on the database focused on types of use and the other worked primarily on the facilitators and barriers to use. Each primary investigator played a secondary role on the other dataset. While working with the data, top-of-mind impressions and memos were written throughout the analytic process to clarify understandings, shape the meaning of codes, and explore links.

Each example of use was given a unique code, identified by source and coded using the conceptual frameworks previously identified. Many examples were initially coded in more than one framework, (e.g., conceptual use in one framework and decision making in another). Following the first review, the second set of interview notes was analyzed. In some cases, new use examples were identified and other use examples were clarified with the coding adjusted accordingly. If discrepancies occurred, the voice recordings were consulted. The growing use database was developed by one of the investigators and exchanged with the other investigator for comment, critique, and challenge to the examples of use and the coding system.

Early on, issues arose with the coding process. First, the definitions of framework codes were not obvious in this context, e.g., what counted as instrumental use in this case? Second, some codes seemed too broad to capture discrete use experiences, such as conceptual use, while other seemed too specific, such as education. Third, the codes began to feel like leaky buckets when working across frameworks. For example, one framework's decision making code spilled into another framework's conceptual use code. Was one of these codes more explanatory to understanding use than the other? Fourth, was the issue of what we called higher order coding.

For example, an identified use might be about decision making, but a related example would show that the decision was turned into action, i.e., someone did something, not just think something. The investigators grappled with the question: *How would the codes account for this link?* Fifth, new codes were needed to capture use examples that didn't fit any of the existing ones. For example, data showed "use" in new ways, with a different audience, beyond the "case" to other times, places and contexts. Sixth, some codes started to break down irreparably. For example, "education of decision makers" failed as a code to capture non-education action with a wide range of stakeholders who make decisions with evaluation findings, including future program participants. These issues around coding occurred not just with frameworks on types of use, but a framework on facilitators and barriers to use. The framework and codes were getting in the way, not showing a way.

Adjustments were made throughout the analytical process; however, a third of the way through coding the interview notes, investigators stopped to review the datasets, the growing problem with codes and the need to address all case study questions. Several key analytical decisions were made. First, we decided to take charge of the codes, rather than abandon them. We defined, adapted, added and relabeled codes as needed to make sense of evaluation use in this case study. Despite problems with the codes, we found they could be a helpful lens and they provided a connection between this case study and the broader evaluation use literature. We did not seek a new language of use; we wanted a clearer understanding of language in context. Second, we coded to more specific meanings or higher order use (action) when a choice was given. For example, if an identified use could be coded as conceptual or decision making, we coded toward the more specific decision making. If however, decision making was tied to action, we coded toward action taken. Third, we added two new case study codes. "Valuing use" emerged as a category since the totality of the evaluation enabled value to be placed on the whole *Active for Life* program, as well as its components, its processes, results and purpose. "Leveraged knowledge" was added to capture the movement of evaluation use across time, contexts and other stakeholders. Fourth, the decision was made to use this revised approach with all interview notes, including those forthcoming and already completed. Lastly, the decision was made to stop coding barriers and facilitators to use into 22 different and overlapping categories and instead tell the story of facilitator of use as it emerged from the data. For example, communication was a leading facilitator mentioned by nearly every informant; no code was needed to identify it as such.

The analysis of evaluation use types continued as an iterative process of identifying, questioning, checking notes, re-classifying and clarifying meaning of codes. Not only did use examples become clearer in the process, but "threads" of evaluation use emerged. These were examples of connections among types of use across time, contexts and categories or codes. Furthermore, the new categories of "valuing use" and "leveraged knowledge" cracked open a new conceptual understanding of evaluation use as ecological. These new understandings were used to tell the story of evaluation use in this case study and organize case study implications.

Study Limitations

There are several limitations to this case study. First, we were unable to transcribe recorded interviews, which would have allowed a greater opportunity to work with and report more of the informants' own words. Second, we were unable to pursue more field and community interviews. While examples of use began to reach redundancy among some stakeholders, it had not with field and community informants when interviews were terminated to meet reporting deadlines. Future research potential exists with these stakeholders. Third, we were not able to pursue "call back" questions due to limitations of time and resources.

Informant Feedback

The case study findings were reviewed and critiqued by the relevant RWJF personnel and sent to all 23 interviewed stakeholders. Such key informants were asked specifically about the plausibility of the use types and facilitators identified and the usefulness of the revised conceptual framework proposed. Such "member checking" adds to the veracity of case study findings (Merriam, 1988).

"This is a nice review of AFL evaluation uses - well documented and cited, theoretically based and supported by the findings."

Anonymous

"I have reviewed the AFL Case Study and I have no changes."

**Melissa Watford, Ed.M.
Health Educator at FirstHealth**

"There were a couple of spots where I have questions—or very minor corrections."

**Diane Dowdy, Ph.D.
NPO Deputy Director, Assistant
Professor at Texas A&M Health
Science Center
Community Health Services**

Active for Life Evaluation Use Study: Tier 1 Questions

1. What was your role in...
 - a. the *Active for Life* program
 - b. the *Active for Life* evaluation
2. What was the purpose of the *Active for Life* evaluation?
3. What were some of the key findings of the *Active for Life* evaluation that could make a difference if *used*?
4. What consideration was given to *Active for Life* evaluation use?
 - a. Was there a specific evaluation *use* plan? When developed? Strategies?
 - b. What would ideal *Active for Life use* look like? What would probable *Active for Life use* look like?
5. Who were the stakeholders of the *Active for Life* evaluation? (review potential list)
 - a. Stakeholder involvement in the evaluation, who? when? how?
 - b. What were facilitators/barriers to stakeholder involvement?
 - c. Of these stakeholders, who were primary or secondary intended users of the evaluation? Non-users?
6. What were the products of the *Active for Life* evaluation? (review list)
 - a. Did these products change over time? If yes, how?
7. What strategies or processes were used to disseminate, communicate, or deliver *Active for Life* evaluation products? (review list)
8. Which products were intended for which users of the evaluation? (how to match?)
9. Are you aware of how findings from *Active for Life* might have been used?
 - a. If yes, examples of use, e.g. who, when, what, how? Corroboration source
 - b. Not sure, what sources might we explore about potential *Active for Life use*
10. In addition to *Active for Life* outcomes, are you aware of any examples of use of the *Active for Life* evaluation process, e.g. who, when, what, how? Corroboration source.
11. Have you personally used the *Active for Life* evaluation findings or process? Example and corroboration
12. What *evidence* do you have about the facilitators or barriers to use of the *Active for Life* evaluation?
13. How could the Robert Wood Johnson Foundation better facilitate the use of its program evaluations?
14. How can the field of [Aging Organizations, Public Health Audience and Medical Community] use the information from the *Active for Life* evaluation?
15. Who else would you suggest we talk with about use of the *Active for Life* evaluation? Where else might we look for evidence of use of the *Active for Life* evaluation?

References

- Alkin MC and Taut SM. "Unbundling Evaluation Use." *Studies in Educational Evaluation*, 29: 1-12, 2003.
- Blake SC and Ottoson JM. "Knowledge Utilization: Implications for Evaluation." In J.M. Ottoson and P Hawe, (eds). Knowledge Utilization, Diffusion, Implementation, Transfer, and Translation: Implications for Evaluation, *New Directions for Evaluation*, No 124. Jossey-Bass: San Francisco, CA: 21-34, Winter 2009.
- Cousins JB and Leithwood KA. "Current Empirical Research on Evaluation Utilization." *Review of Educational Research*, 56(3): 331-364, 1986.
- Green LW, Potvin L and Richard L. "Ecological Foundations of Health Promotion." *American Journal of Health Promotion*, 10(4): 270-281, 1996.
- Henry GT and Mark MM. "Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions." *American Journal of Evaluation*, 24(3): 293-314, 2003.
- Johnson K, Greenside LO, Toal SA, King JA, Lawrenz F and Volkov B. "Research on Evaluation Use: A Review of the Empirical Literature from 1986 to 2005." *American Journal of Evaluation*, 30(3): 377-410, 2009.
- King JA and Pechman EM. "Pinning the Wave to the Shore: Conceptualizing Evaluation Use in School Systems." *Educational Evaluation and Policy Analysis*, 6(3): 241-253, 1984.
- Kirkhart KE. "Reconceptualizing Evaluation Use: An Integrated Theory of Influence." *New Directions for Evaluation*, 88: 5-24, 2000.
- Leviton LC and Hughes EFX. "Research on the Utilization of Evaluations: A Review and Synthesis." *Evaluation Review*, 5(4): 525-548, 1981.
- Leviton LC and Boruch RF. "Contributions of Evaluation to Education Programs and Policy." *Evaluation Review*, 7: 563-598, 1983.
- Merriam SB. *Case Study Research in Education: A Qualitative Approach*. San Francisco, CA: Jossey-Bass, 1988.
- Merriam SB. *Qualitative Research: A Guide to Design and Implementation*. San Francisco, CA: Jossey-Bass, 2009.

Ottoson JM, Rivera MD, DeGroff A, Hackley S and Clark C. On the Road to the National Objectives: A Case Study of Diabetes Prevention and Control Programs. *Journal of Public Health Management and Practice*, 13(3):287-295, 2007.

Ottoson JM and Hawe P. Knowledge Utilization, Diffusion, Implementation, Transfer, and Translation: Implications for Evaluation, *New Directions for Evaluation*, No 124. San Francisco, CA: Jossey-Bass, 2009.

Patton MQ. *Utilization-Focused Evaluation*. Thousand Oaks, CA: Sage, 1997.

Pressman J L and Wildavsky A. *Implementation, 3rd ed*. Berkeley, CA.: University of California Press, 1984.

Rich RF. "Uses of Social Science Information by Federal Bureaucrats: Knowledge for Action Versus Knowledge for Understanding," in CH Weiss (ed.) *Using Social Research in Public Policy Making*. Lexington, MA: Lexington Books, 1977.

Shadish WR, Cook TD and Leviton LD. *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: Sage Publications, 1991.

Stake RE. *The Art of Case Study Research*. Thousand Oaks, CA: Sage, 1995.

Taut SM and Alkin MC. "Program Staff Perceptions of Barriers to Evaluation Implementation." *American Journal of Evaluation*, 24(2): 213-226, 2003.

The Joint Committee on Standards for Educational Evaluation. *The Program Evaluation Standards, 2nd ed*. Thousand Oaks, CA: Sage, 1994.

Weiss CH. "The Many Meanings of Research Utilization." *Public Administration Review*, September/October: 426-431, 1979.

Weiss CH. "Knowledge Creep and Decision Accretion." *Knowledge: Creation, Diffusion, Utilization*, 1(3): 381-404, 1980.

Yin RK. *Case Study Research: Design and Methods, 3rd ed*. Thousand Oaks, CA: Sage, 2003a.

Yin RK. *Applications of Case Study Research, 2nd ed*. Thousand Oaks, CA: Sage, 2003b.

Bibliography

Publications of Active Choices and Active Living Every Day Evidence-Based Trials

Dunn AL, Marcus BH, Kampert JB, Garcia ME, Kohl HW and Blair SN “Comparison of Lifestyle and Structured Interventions to Increases Physical Activity and Cardiorespiratory Fitness: A Randomized Trial.” *JAMA*, 281(4): 327-334, 1999.

King AC, Haskell WL, Taylor CB, Kraener HC and DeBusk RF. “Group-Versus Home-Based Exercise Training in Healthy Older Men and Women: A Community-Based Clinical Trial.” *Journal of the American Medical Association*, 266: 1535-1542, 1991.

King AC and Brassington G. “Enhancing Physical and Psychological Functioning in Older Family Caregivers: The Role of Regular Physical Activity.” *Annals of Behavioral Medicine*, 19: 91-100, 1997.

King AC, Pruitt LA, Phillips W, Oka R, Rodenburg A and Haskell WL. Comparative Effects of Two Physical Activity Programs on Measured and Perceived Physical Functioning and Other Health-Related Quality of Life Outcomes in Older Adults.” *Journal of Gerontology: Medical Sciences*, 55A: M74-M83, 2000.

King AC, Baumann K, O’Sullivan P, Wilcox S and Castro C. “Effects of Moderate-Intensity Exercise on Physiological, Behavioral, and Emotional Responses to Family Caregiving: A Randomized Controlled Trial.” *Journal of Gerontology: Medical Sciences*, 57A: M26-M36, 2002.

King AC, Friedman RH, Marcus B, Castro C, Napolitano M, Ahn D and Baker L. “Ongoing Physical Activity Advice by Human Versus Computers: The Community Health Advice by Telephone (CHAT) Trial.” *Health Psychology*, 26(6):718-727, 2007.

Publications with the Results from the *Active for Life* Program

Griffin SF, Wilcox A, Ory MG, Lattimore D, Leviton L, Castro C, Carpenter RA and Rheaume C. "Results from the Active for Life: Process Evaluation: Program Delivery Fidelity and Adaptations." *Health Education Research*, 25(2): 325-342, 2010.

Wilcox S, Dowda M, Griffin SF, Rheaume C, Ory MG, Leviton L, King AC, Dunn A, Buchner DM, Bazzarre T, Estabrooks PA, Campbell-Voytal K, Bartlett-Prescott J, Dowdy D, Castro CM, Carpenter, RA, Dzewaltowski DA, Mockenhaupt R. "Results of the First Year of Active for Life: Translation of 2 Evidence-Based Physical Activity Programs for Older Adults into Community Settings." *American Journal of Public Health*, 96(7): 1201-1209, 2006.

Wilcox S, Dowda M, Leviton LC, Bartlett-Prescott J, Bazzarre T, Campbell-Voytal K, Carpenter RA, Castro CM, Dowdy D, Dunn AL, Griffin SF, Guerra M, King AC, Ory MG, Rheaume C, Tobnick J, Wegley S. "Active for Life: Final Results from the Translation of Two Physical Activity Programs." *American Journal of Preventive Medicine*, 35(4): 340-351, 2008.

Wilcox S, Dowda M, Wegley S and Ory MG. "Maintenance of Change in the Active-for-Life Initiative." *American Journal of Preventive Medicine*, 37(6): 501-504, 2009.