Yates, B. T. (2012). Step arounds for common pitfalls when valuing resources used versus resources produced. In G. Julnes (Ed.), *Promoting valuation in the public interest: Informing policies for judging value in evaluation. New Directions for Evaluation, 133,* 43–52.

4

Step Arounds for Common Pitfalls When Valuing Resources Used Versus Resources Produced

Brian T. Yates

Abstract

The value of a program can be understood as referring not only to outcomes, but also to how those outcomes compare to the types and amounts of resources expended to produce the outcomes. Major potential mistakes and biases in assessing the worth of resources consumed, as well as the value of outcomes produced, are explored. Most of these occur when the evaluation is limited in contexts examined or perspectives adopted. In particular, it is noted that the price of a resource often is context-dependent, and may not describe the value of a resource from important perspectives. Also, the monetary value of outcomes as inferred from earned income, and from avoided human service expenditures, may not reflect outcome value from key perspectives, possibly exacerbating discrimination according to gender, ethnicity, and age. Solutions for these problems are recommended. More complete and detailed information on resources consumed and outcomes produced also may facilitate systematic optimization of program value, if the evaluation includes the amounts and types of resources used by those program activities that change the participant processes that lead to desired program outcomes. © Wiley Periodicals, Inc., and the American Evaluation Association.

> The cynic . . . knows the price of everything and the value of nothing.
>
> —Oscar Wilde's *Lady Windermere's Fan*

NEW DIRECTIONS FOR EVALUATION, no. 133, Spring 2012 © Wiley Periodicals, Inc., and the American Evaluation Association. Published online in Wiley Online Library (wileyonlinelibrary.com) • DOI: 10.1002/ev.20005

Evaluators rarely are cynics . . . at least at the start of an evaluation. Also of note about evaluators is that they rarely consider the types and amounts of resources consumed by a program when evaluating it. Instead, the activities of the program and the reactions of participants and onlookers to those activities often are the primary focus. Even evaluations that appear to include costs typically report only the summed current value of resources purchased for use in the program, that is, the total price paid, which often conveys less information than is needed to describe programs accurately and to optimize programs systematically.

For example, how students react emotionally to the educational process in which they have participated is captured qualitatively and measured quantitatively in many evaluations of teaching. Changes in students' knowledge of facts and theories, and students' performance of skills from computations to therapies, also are assessed in some of evaluations of teaching. Even the amount of money spent per student may be reported. But does any of this help us judge the *value* of the education that was received? Not really. Furthermore, does this evaluation of teaching help us understand what actually happened, why some goals were achieved and others were not, and most importantly how to *improve* the education? Again, no.

Knowing the summary price at which a service was purchased does not establish the value of the service, especially when that service is not available in a free market and purchasable by a large number of rational decision-makers (as is common in many programs). Instead, we can begin the valuing process of most human services by asking what made the service possible. For teaching, that would be the training, time, and effort of the teachers, most likely. Anything else? Perhaps the time spent or not spent by students and possibly their parents or tutors in educational activities, plus the knowledge and skills students already have or have not acquired, could contribute to or detract from education outcomes. Are any other resources used in a program of education? Brick-and-mortar buildings, possibly . . . and some form of information technology, certainly, whether texts or computers, probably supported by administrative and other services.

These resources that make possible program operations and outcomes are ignored entirely or mentioned tangentially in most evaluations, despite decades of efforts to persuade evaluators to include costs in evaluations of education, mental health, substance-abuse treatment, and other programs (e.g., Carter & Newman, 1976; Fishman, 1975; Levine & McEwan, 2001; Sorensen & Phipps, 1975; Yates, 1980, 1996). Many evaluations compare outcomes of these programs. Most evaluators probably recognize that different amounts or types of resources could produce different outcomes for the programs they evaluate . . . yet exceptionally few evaluators and evaluations report or compare costs!

Ignoring the *contextual pragmatics* of those resources available to, and used by, a program risks making attributions about the potential effectiveness

of programs that may be more accurately attributed to the resources that were or were not available at the sites at which the program was implemented. Consider, for instance, the inaccuracy as well as the inequity of comparing educational outcomes at well- and poorly funded schools (cf. Ross, Barkaoui, & Scott, 2007). Blaming relatively poor student performance on the educational methods, teachers, principals, or students could be incorrect, and could lead to harmful personnel and funding decisions, if the school had been underfunded relative to other schools. Yet most evaluations exclude meaningful information on the types and amounts of resources that were and were not available to the program in the context in which it operated during the evaluation. Particularly in times of increasing constraints on many resources needed by human services, program evaluations seem incomplete and even impractical unless they include the resources available to and used by programs in the contexts in which they try to operate. Particularly because outcomes are distal and probabilistic while resource expenditures are more proximal and certain, including in an evaluation information about the resources consumed by a program seems at least as important as information about outcomes attained by a program.

Step Arounds for Pitfalls and Biases in Evaluating Resources Consumed by Programs

Evaluating program costs involves more than simply listing *prices* paid at a particular time and place to acquire the resources used by the program. The money one had to pay to obtain these resources can be a poor representation of their worth, as Mr. Wilde noted at the start of this manuscript. Price is a function of contextual features such as demand by other programs and availability, and can be manipulated to create artificial scarcity or exaggerate apparent abundance. Rather, a cost-inclusive evaluation can better help a program achieve its goals by including a description of the types and amounts of specific resources used by a program, plus the specific activities that those resources were used to enable (cf. Yates, 1996). This information allows managers to adjust the degree to which different program activities are implemented in response to changes in the resources available (cf. Yates, 1980).

Substantial errors in managing, understanding, and disseminating programs in different contexts can be introduced if an evaluation provides an incomplete description of resources used by a program at particular sites. These incomplete resource valuations are more common in mono-perspective evaluations, in addition to other problems described by Alkin, Vo, and Christie (this issue). Too often, evaluations that attempt to incorporate costs ignore resources contributed by program participants, such as time spent by clients when receiving program services. This *nonvaluing* of participant time can be interpreted as profoundly *devaluing* the client, a bias that can lead to

excessive reliance on resources that are scarce for many clients. Client time spent receiving services has intrinsic value, to clients and client advocates at least. Client time may have no monetary value or "price" according to some stakeholders, in that the client may not be employable or may not work outside the home. The cost to family members of replacing some clients' time in the home can be considerable,however, if childcare or housework was being provided by the client prior to inpatiency. Similarly, unless the client resides in the program facility, time and expense traveling to and from program sites is an additional resource seldom included in program evaluations. Excluding these and other client resource expenditures not only provides an incomplete depiction of the program, but also underestimates the value of total resources consumed by the program.

Moreover, clients of most human services are asked to seek services in addition to those of the program being evaluated, such as vocational training or support groups. Client time, client funds, and client transportation spent because of these referrals, and the resources consumed by the referred-to services when actually used, can be critical to include. Otherwise an evaluation might erroneously report substantial returns for minimal investments that actually required more resources of society because of unreported use of other programs' services.

In addition, time spent by service providers may be underrepresented by time recorded in payroll records. In primary and second school, for example, teachers may spent entire evenings beyond their recognized working hours grading and commenting on assignments and tests. In mental health services, too, some providers receive little or no money, volunteering substantial time and costly transportation resources in exchange for supervision and training at externships and internships required for graduate degrees and licensure (e.g., Yates, Haven, & Thoresen, 1979).

Resources other than provider and client time may be consumed but not reported accurately or at all in budgets. These include donated facilities, equipment and materials donated, and administrative services and utilities paid for by a parent organization but omitted from the evaluation report (e.g., Yates et al., 2011). In sum, a full accounting of resources used by a program describes what the program uses and consumes. Accounting records alone can underrepresent those resources in ways that can distort evaluation findings.

Evaluating Program Value as What Is Produced Relative to What Is Consumed

> When judging the value of a program in the context of constrained funding, outcomes are meaningless until they are compared to the types and amounts of resources consumed to produce them.
>
> —Anonymous

Where does such a bold statement come from? From the perspective of many decision makers who use evaluations! For funders, programs are means of transforming those resources that are available in a community into something the funder is responsible to foster or maintain—turning uneducated children into educated adults, for example. Moreover, in light of the diminishing availability of many resources for most needs, funders also weigh not only whether more resources were produced than were consumed by individual programs, but how the net benefit of some programs compares to the net benefit of other, alternative programs. For many programs, including substance-abuse treatment and psychological services, benefits such as increased licit income and reduced use of health, mental health, and criminal justice services are 9.7 to 14.9 times larger than the costs of those programs—see, for example, French et al. (2000).

So, is this what *value*-ation of a program has come to: comparison of the resources used (costs) versus the outcomes produced (benefits)? Yes. Must we evaluators really consider doing cost–benefit analysis now? Could not this attention to program costs, benefits, and net benefit be delegated to other professionals, or postponed until we have thoroughly evaluated the outcomes of a program?

Yes, indeed. We evaluators do not have to evaluate costs and benefits and compare the two. Others will, and increasingly are. *Their* findings are the ones that will decide which programs receive public monies and which will not. In contrast to the tentative approaches of the legislatures of the United States (Shipman, this issue) and Canada (Dumaine, this issue), some state legislatures have been using available cost–benefit analyses to decide which programs to fund. Building on initial findings from the turn of the century, for example, the Washington State legislature in 2009 directed its policy institute to: ". . . calculate the return on investment to taxpayers from evidence-based prevention and intervention programs and policies." The legislature instructed the institute to produce "a comprehensive list of programs and policies that improve . . . outcomes for children and adults in Washington and result in more cost-efficient use of public resources" (Aos et al., 2011, p. 1).

The results of funding only those programs that showed the highest return on investment of taxpayer funds already have convinced the Washington state legislature to continue making funding decisions according to findings of evaluations that include monetary benefits as well as costs:

> Today, the results of these crime-focused efforts appear to be paying off. Relative to national rates, juvenile crime has dropped in Washington, adult criminal recidivism has declined, total crime is down, and taxpayer criminal justice costs are lower than alternative strategies would have required. (Aos et al., 2011, page 1)

My advice: we evaluators should abandon our apparent *value phobia* (Scriven, this issue) with regard to cost–benefit relationships as well as

valuation more generally. We should embrace and report multiperspective valuations of those resources that make possible the specific activities that lead to targeted program outcomes.

Step Arounds for Pitfalls and Biases in Evaluating Resources *Produced* by Programs

To minimize exaggeration of program benefits relative to program costs, resources produced by a program need to be evaluated at least as thoroughly and carefully as the resources consumed by the program (cf. Glick, Doshi, Sonnad, & Polsky, 2007; Gold, Siegel, Russell, & Weinstein, 1996). However, just as boiling down information on diverse resources consumed by a program into a single "price" can hide omission of resources critical to understanding and establishing a similar program in another setting, so can reducing the many outcomes of a program into a single monetary figure exclude information essential to an accurate understanding of what a program accomplishes. One common outcome measured in evaluations of resources generated by programs is hoped-for increments in client income. Another outcome commonly measured in benefit valuations is savings that can result from participation in program activities, such as when substance-abuse treatment reduces subsequent use of health and criminal justice services (e.g., Mannix, 2010). Both income enhancement and service savings typically are valued in local currency units, such as additional dollars earned and dollars saved. Often these are summed to form a measure of the *benefits* of the program. Distilling these resource-related outcomes into a single number can hide errors in valuation.

For example, clients whose incomes differ not according to productivity but according to gender, age, ethnicity, or country may appear to benefit to different extents from the same program. A program that put back to work a 35-year-old male could appear, for instance, to be more beneficial in future annual income earned than a program that returned to work a 35-year-old female, due to ignominious disparities in pay for males versus females. Also, in terms of projected total earnings over one's lifetime, a program that returned to work a 35-year-old female could appear more beneficial than one that returned to work a 65-year-old male, because of the likely longer remaining lifetimes of the female versus the male. Furthermore, a substance-abuse treatment program that reduces health care utilization for males might appear less beneficial than a substance-abuse treatment program that reduces the higher levels of health care utilization found for most females, even though the same proportional reduction in service utilization was produced for both genders (cf. Mannix, 2010). Finally, treatment and prevention of a disease that diminishes productivity could be viewed as more beneficial for countries in which residents had higher earning potential.

One possible solution to these problems is to accept the observed or projected differences in income and service savings as valid indicators of differential program benefits. It has been argued that this practice could lead to funding decisions for employment programs that perpetuate or exacerbate current social inequities for women versus men, for different ethnicities, and for the aging (Yates, 1986). A preferable alternative may be to use pay rates that are standardized across genders, ethnicities, ages, countries, and other variables on which differences in income and health care use exist. Also, constants reflecting the complete value of a health service could be used to monetize future expenditures for those more accurately, rather than using the rates set for publicly funded services, such as Medicaid and Medicare. The latter could underestimate the benefits of services, such as substance abuse treatment, that reduce future health care expenditures.

Conducting evaluations in ways that consider these potential problems in measuring program outcomes could provide more socially valid assessments of program value, as called for by Morris (this issue). An alternative approach to attempting to monetize program outcomes is to find a unit of outcome measurement that can be combined with measures of program resources while being sufficiently general to allow comparison of a variety of programs. One promising measure is the quality-adjusted life year or QALY (cf. Drummond, Sculpher, Torrance, O'Brien, & Stoddard, 2005; Gold et al., 1996; e.g., Freed, Rohan, & Yates, 2007). QALYs have been standardized for a variety of medical and psychological treatments via surveys that ask individuals to judge what proportion of a year of complete health would be equivalent to a full year of life with a particular malady. Combining this information with findings on program effectiveness allows expression of a health improvement as a portion of a QALY gained due to the program. The relative value of programs then can be compared in terms of QALYs gained. The societal value of these programs is assessed by including findings on the value of the resources that were consumed to cause the increment in QALYs, as in "cost per quality-adjusted life year gained." This also is one major way to include measures of outcomes that, as Morris (this issue) notes, are more important to some interest groups than increments in productivity and decrements in future spending for other services.

The Biggest Step Around: Evaluate *More* Than Resources Consumed and Resources Produced

Questions about the value of a program can stop at the summative, "Does the program generate more, or fewer, resources than it consumes?" or the comparative summative, "Does the program produce more resources relative to those it consumes than is the case for other programs?" These are important questions that, when answered in ways that avoid the pitfalls explored above, can help decision makers better allocate societal resources for the

collective good. If evaluation has a more *formative* function, however (Chelimsky, 1997; Scriven, 1967), it may be helpful to understand what did and did not occur between the consumption of resources and the hoped-for generation and savings of resources. This understanding is aided, and may only be possible, if the evaluation includes information on the types and amounts of activities in which participants engage. Even this third type of information may not be enough. If the typical causal chain of events is hypothesized, a complete theory or model of a program would include:

- Resources consumed by the program
- Activities made possible by those resources, and that the program used to engage participants
- Processes that occurred inside the brains and bodies of participants as a result of program activities
- Outcomes that were observed as a result of those processes (Yates, 1996)

This resource–activity–process–outcome (RAPO) framework can help evaluators model and understand how a program produces its outcomes. This sort of more complete model also can help an evaluator understand why a program yielded outcomes that were the opposite of those it was designed and funded to achieve, for example, increased rather than decreased use of alcohol, tobacco, and other drugs. For instance, RAPO analysis of a drug-use prevention program indicated that program activities that consumed the fewest resources per student (small student groups, which used little teacher and parent time, and little classroom space) actually decreased rather than increased social responsibility in fourth-grade students (cf. Yates, 2002). This decreased social responsibility, was, in turn, associated with higher use of alcohol, tobacco, and other drugs. If individual student participation in each of the specific activities of this substance-use prevention program had not been measured, and several psychosocial processes (including social responsibility) had not been assessed as well, the specific iatrogenic component of the program might not have been identified in the evaluation.

An additional advantage of including activities and processes as well as resources and outcomes in an evaluation is the possibility that these four variables could be combined into a quantitative model of the program. Such a model can be manipulated to find ways to optimize relationships between resources used and outcomes achieved by the program (Yates, 1996). After entering resource, activity, process, and outcome data, the model manipulation technique *linear programming* finds the combination of activities that should maximize program outcomes within the specific resource constraints of the program context (Yates, 1980). Similar analyses could find the mixture of activities that should minimize the costs of achieving specific outcomes. Maximization of benefits given available resources, and minimization of resources consumed to achieve set benefits, are especially compelling visions of how program value might not only be evaluated, but optimized.

Conclusions

1. Evaluations that exclude information on the types and amounts of resources available and used by programs may arrive at erroneous conclusions that could lead to harmful funding decisions.
2. Prices paid for resources used by programs may not provide accurate, useful descriptions of those resources.
3. Valuing resources used by programs from multiple perspectives, such as those of clients and providers, can provide more complete and accurate, and less biased, cost assessment.
4. Multiple perspectives, such as those of clients, decision makers, and taxpayers, also can describe and value the types and amounts of resources *produced* by programs in a more accurate and less biased manner.
5. Program outcomes need not be measured as money produced or saved, even when comparing those outcomes to program costs. Outcomes of diverse human services can be expressed as QALY outcomes, for example, and can be compared in terms of resources consumed to produce those QALYs.
6. Program value can be optimized if the evaluation includes information on activities and resulting changes in client processes that contribute to program outcomes, in addition to the resources used and outcomes yielded.
7. Given the above points, comparison of program outcomes generated to program resources used, or cost–benefit analysis, need not be a narrow lens on program performance. Multiperspective RAPO analyses can provide a comprehensive, formative, and influential means of providing the best services to the most people for the least amount of resources.

References

Aos, S., Lee, S., Drake, E., Pennucci, A., Klima, T., Miller, M., . . . & Burley, M. (2011). *Return on investment: Evidence-based options to improve statewide outcomes* (Document No. 11–07–1201). Olympia, WA: Washington State Institute for Public Policy.

Carter, D. E., & Newman, F. L. (1976). *A client-oriented system of mental health service delivery and program management: A workbook and guide* (pp. 76–307). Rockville, MD: Department of Health, Education, & Welfare.

Chelimsky, E. (1997). The coming transformation in evaluation. In E. Chelimsky & W. R. Shadish (Eds.), *Evaluation for the 21st century: A handbook.* Thousand Oaks, CA: Sage.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddard, G. L. (2005). *Methods for economic evaluation of health care programs.* Oxford: Oxford University Press.

Fishman, D. B. (1975). Development of a generic cost-effectiveness methodology for evaluating patient services of a community mental health center. In J. Zusman & C. R. Wurster (Eds.), *Program evaluation: Alcohol, drug abuse, and mental health services.* Lexington, MA: Lexington Books.

Freed, M. C., Rohan, K. J., & Yates, B. T. (2007). Estimating health utilities and quality adjusted life years in seasonal affective disorder research. *Journal of Affective Disorders, 100,* 83–89.

French, M. T., Salomé, H. J., Krupski, A., McKay, J. R., Donovan, D. M., Mclellan, A. T., & Durell, J. (2000). Benefit–cost analysis of residential and outpatient addiction treatment in the State of Washington. *Evaluation Review, 24,* 609–634. DOI: 10.1177/0193841X0002400603

Glick, H. A., Doshi, J. A., Sonnad, S. S., & Polsky, D. (2007). *Economic evaluation in clinical trials.* London: Oxford University Press.

Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (Eds.). (1996). *Cost-effectiveness in health and medicine.* New York: Oxford University Press.

Levine, H. M., & McEwan, P. J. (2001). *Cost-effectiveness analysis* (2nd ed.). Thousand Oaks, CA: Sage.

Mannix, D. (2010). *Gender differences in the costs, benefits, cost-effectiveness, and cost-benefit of substance abuse treatment: A study in four treatment settings* (Unpublished doctoral dissertation). Department of Psychology, American University, Washington, DC.

Ross, J. A., Barkaoui, K., & Scott, G. (2007). Evaluations that consider the cost of educational programs: The contribution of high-quality studies. *American Journal of Evaluation, 28,* 477–492.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation.* Chicago, IL: Rand-McNally.

Sorensen, J. E., & Phipps, D. W. (1975). *Cost-finding and rate-setting for community mental health centers.* Rockville, MD: Department of Health, Education, & Welfare.

Yates, B. T. (1980). *Improving effectiveness and reducing costs in mental health.* Springfield, IL: Thomas.

Yates, B. T. (1986). Economics of suicide: Toward cost-effectiveness and cost-benefit analysis of suicide prevention. In R. Cross (Ed.), *Non-natural death: Coming to terms with suicide, euthanasia, withholding or withdrawing treatment.* Denver, CO: Rose Medical Center.

Yates, B. T. (1996). *Analyzing costs, procedures, processes, and outcomes in human services: An introduction.* Thousand Oaks, CA: Sage Publications.

Yates, B. T. (2002). Roles for psychological procedures, and psychological processes, in cost-offset research: Cost → procedure → process → outcome analysis. In N. A. Cummings, W. T. O'Donohue, & K. E. Ferguson (Eds.), *The impact of medical cost offset on practice and research: Making it work for you* (pp. 91–123). Reno, NV: Context Press.

Yates, B. T., Haven, W. G., & Thoresen, C. E. (1979). Cost-effectiveness analysis at Learning House: How much change for how much money? In J. S. Stumphauzer (Ed.), *Progress in behavior therapy with delinquents* (pp. 186–222). Springfield, IL: Charles C. Thomas.

Yates, B. T., Mannix, D., Freed, M. C., Campbell, J., Johnsen, M., Jones, K., & Blyler, C. (2011). Consumer-operated service programs: Monetary and donated costs and cost-effectiveness. *Psychiatric Rehabilitation Journal, 35*(2), 91–99.

BRIAN T. YATES is a professor in the Department of Psychology at the American University in Washington, DC, where he teaches, conducts research on program evaluation, and directs the Program Evaluation Research Lab (PERL). Most of his publications apply cost-effectiveness or cost-benefit analysis to the systematic evaluation and improvement of human services.