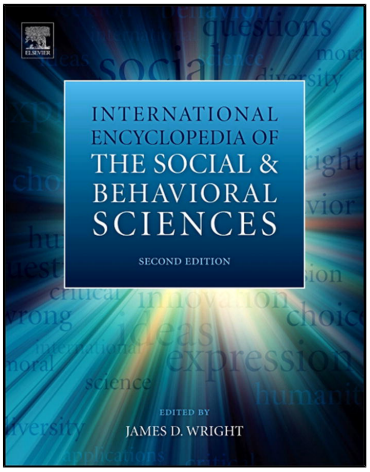

From Yates, B.T., 2015. Cost-Benefit and Cost-Effectiveness Analyses in Evaluation Research. In: James D. Wright (editor-in-chief), International Encyclopedia of the Social & Behavioral Sciences, 2nd edition, Vol 5. Oxford: Elsevier. pp. 55–62.
ISBN: 9780080970868

Cost-Benefit and Cost-Effectiveness Analyses in Evaluation Research

Brian T Yates, Department of Psychology, American University, Washington, DC, USA

Abstract

Cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA) are used increasingly in both clinical trials of health and human services, and in evaluations of programs to determine their future funding. Definitions of CBA and CEA are given and contrasted with each other and with requests for 'cost analyses' and 'efficiency' studies. The emerging importance of cost-utility analysis is noted as well. Emerging directions in cost-inclusive assessment are noted, including evaluating costs of using different services to add Quality-Adjusted Life Years and Disability-Adjusted Life Years – nonmonetary measures that can be almost as useful for comparing outcomes of diverse programs as monetary measures, i.e., program benefits. Advantages of conceptualizing 'costs' more broadly as the value of the specific types and amounts of resources used to provide services are explained. Ethical problems in cost-inclusive evaluation are highlighted, and possible solutions suggested. Viewing CBA and CEA as crucial parts of improvement-oriented operations research is recommended. Links for manuals and Web sites for cost-inclusive evaluation in health and human services are provided.

Definitions

Cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA) are two related methods for social scientists and program evaluators to incorporate data on resource requirements of, and resources generated and saved by, social services. At their most basic, both forms of cost-inclusive evaluation compare the value of resources invested in an activity to the value of the outcomes of that activity (Yates, 2009). The activity being examined for cost-benefit or cost-effectiveness can be anything from introducing free health services in a remote village to global eradication of an illness. Persons and organizations can use findings from CBA and CEA to decide whether to engage in or continue an activity and how much time, money, and effort to devote to it. Using different stakeholder perspectives to define, measure, and compare costs, benefits, and effectiveness can result in different findings, different funding, and different impacts for individuals, communities, and human society.

As resource constraints increase, a health or human service tends to be funded only if total benefits of that service exceed total costs of that service by a substantial amount – and relatively soon. Conceptualizing services as human rights or entitlements can drive decision-makers to select the least costly activity, paying less attention to outcomes. This can be unwise: social scientists have developed reliable and valid means of measuring most program outcomes, and of comparing the monetary value of those outcomes over time to program costs over time.

Program Evaluation

Answering questions about how faithfully a program is being implemented, whether its goals are being achieved, and whether the resources devoted to program operations are being used in the most worthwhile manner, all can be considered forms of program evaluation (cf Posavac, 2011). Many research strategies can be used to answer these questions, ranging from expensive and possibly invasive randomized clinical trials to often less expensive but sometimes less valid quasiexperiments (American Evaluation Association, 2013; Shadish et al., 2001). Program evaluation often involves experts from diverse disciplines who apply quantitative and qualitative means of judging programs as they currently operate (*summative* program evaluation), and of improving programs so they operate better (*formative* program evaluation; cf Yates, 1996).

Goals, Effectiveness, Efficacy, and Efficiency

Evaluating the degree to which an activity's desired outcomes or *goals* are being achieved is a judgment of the *effectiveness* of the activity. Activities shown by research to be promising are bundled into programs designed to prevent or remediate problems. Researchers monitor the *fidelity* with which each activity is implemented by highly trained program provider. Persons receiving program services (*clients*) may be assigned randomly to either the program or various control conditions to allow effects of other events, of measurement itself, and of client expectations to be separated from the unique contributions of the program. Participants also may be screened so they have similar characteristics – sometimes so they have only the problem the program is designed to remediate, such as depression, and no other comorbidities, such as anxiety or diabetes. Under these controlled and, some would say, unrealistic conditions, the maximum effectiveness or *efficacy* of the program is measured. When implemented later by providers with varied training and more demands on their time, for clients with combinations of problems, and in programs that must not consume more resources than they will be reimbursed, the *effectiveness* of the program in the 'real-world' is evaluated. Often *effectiveness* is less than *efficacy*.

The 'efficiency' of a program can be distinguished from both its effectiveness and efficacy. In some contexts, 'efficiency' refers to how well the program uses those resources currently available in its environment. This usage resembles *cost-effectiveness* and *cost-benefit* (e.g., Marseille et al., 2004). Other meanings of 'efficiency' may include simple cost per student or cost per patient. Again, 'efficiency' does not have one widely accepted definition, although it appears in speeches about program costs about as often as, and sometimes interchangeably with, 'cost-benefit' and 'cost-effectiveness.' One even hears 'cost-efficiency.'

International Encyclopedia of the Social & Behavioral Sciences, Second Edition, 2015, 55–62

Actually, *cost*, *cost-benefit*, and *cost-effectiveness* all are terms with specific definitions.

Cost Analysis

An entreaty for 'cost-efficiency,' 'cost-benefit analysis,' or 'cost-effectiveness analysis' often is a request for greater attention to program costs and not a thorough analysis of relationships between costs and benefits or costs and effectiveness. In program evaluation, *costs* usually refer to the total monetary value of the many resources consumed by program activities. A *cost analysis* compares programs purely on their costs, often assuming that any differences in program benefits or effectiveness are minor or unmeasurable. If programs provide entitlements such as income support, aid for children, or health services, program outcomes may not even be measured. Evaluating programs according to their costs but not their outcomes does, of course, risk funding programs that cost the least but also deliver the least, to the extent that program outcomes are functions of the types and amounts of resources available to the program. It also is possible that more-costly programs deliver poorer outcomes than less-costly programs (e.g., Yates, 1994).

Usually *cost analysis* focuses on only the *monetary* value of activities performed by providers. Comprehensive cost analysis performed from the perspective of society considers donated resources such as time volunteered by interns and underpaid providers, and clients.

Benefit Analysis

In cost-inclusive evaluation, *benefits* do not refer to positive outcomes in general, but only monetary outcomes. Most providers of health and human services are trained to focus on changing key client processes and proximal outcomes rather than more distal outcomes. For example, a psychologist may implement activities within cognitive-behavioral therapy and attend primarily to the immediate results of those activities, such as whether the problematic thought patterns changed and anxiety diminished. Outcomes of greater interest to health insurers, employers, and family members might be reduced use of health services due to decreased need, less self-medication with alcohol and other drugs, increased days worked, improved productivity or profitability, and greater availability to provide child- or elder care. These outcomes are typically what programs are funded to achieve. They are either monetary or readily monetizable using methods detailed in monographs (e.g., Gold et al., 1996; Drummond et al., 2005; Yates, 1980, 1996, 1999). All are termed *benefits* and are summed to describe program outcomes to third parties, i.e., to entities other than the client and provider. These can include *negative benefits* such as increased use of health and other human services following treatment, which can be viewed as undesirable by third parties but as ideal by practitioners, clients, and consumer advocates.

Cost-Benefit Analysis

Once summed, the monetary value of activity outcomes can be compared to the total monetary value of resources contributed by providers, clients, and others because those outcomes and resources were measured in the same units. The resulting difference (benefits minus costs or *net benefit*) and ratio (benefits/costs) seem easy to understood and use when choosing among programs to fund (e.g., Yeh, 2009). Programs may differ in whether they are cost-beneficial or not, i.e., whether their net benefit exceeds 0 or their benefit/cost ratio exceeds 1. Programs also can differ in how *much* their benefits exceed costs, either in absolute or relative values, and how soon this can be expected to occur. Because 'benefits' precede 'costs' when describing formulas for both net benefit and benefit-cost ratios, cost-benefit analysis (CBA) might be more understandably termed *benefit-cost analysis* or BCA. The benefit can be seen as a *return* on investment, social or individual, and the amount and speed of that return relative to the value of resources spent is the *return on investment* or ROI. Returns on investment that occur several decades after resource expenditure, as is common in prevention and education programs, are not as valuable as the same returns occurring sooner. This difference is routinely quantified by adjusting temporally distant returns downward for their *net present value* using a *discount rate* to express the return that would have been achieved if the same resources had been invested in alternative activities. Several discount rates are used, resulting in not one but several 'bottom lines' when time to ROI (TROI) and total amount of ROI are reported.

Neglecting certain costs or ignoring particular benefits can, of course, alter indices of cost-benefit to support decisions that later prove to be incorrect. Also, some outcomes are difficult to convert into monetary units, especially in a manner that does not allow under- or overrepresentation due to subjective values of the evaluator or decision-maker. Moreover, some services are considered rights to which all persons are entitled regardless of whether they are or are not found to have benefits which exceed costs. These entitlements include, in many areas of our world, healthy air and water, adequate and nutritious food, opportunities for physical activity, reliable shelter, child- and elder care, protection from physical harm and property theft, comprehensive health care, access to information resources as provided by libraries and the Internet, universal education, opportunities for fulfilling employment, income maintenance when one is not employed, liberty, and happiness (cf United Nations, 2014).

Cost-Effectiveness Analysis (CEA)

When nonmonetary and nonmonetized outcomes are compared to monetary costs for a single program, simple differences or ratios do not yield the same simple decision aids as net benefit or benefit-cost ratios because outcomes are measured in different units than costs. Calculating the 'cost per' unit of nonmonetary outcome may not seem sensical at first unless two or more programs are being compared, but consider the potential usefulness of knowing the *cost per drug-free month* when comparing substance abuse treatment programs, or even the *cost per pound lost* when comparing weight-loss regimens. This is different than 'cost per patient-filled bed,' unless simply filling a bed is the outcome of interest. To be clear: 'Cost per student' would be a measure of cost only. 'Cost per student who graduated with scores exceeding minima required' would be a measure of cost-effectiveness.

Cost-Utility Analysis

An increasingly common form of cost-effectiveness analysis (CEA), *cost-utility analysis (CUA)*, uses nonmonetary measures of outcome that can be applied to many programs. Essentially, program outcomes are measured as not program-specific measures such as 'apples' or 'oranges,' but as more general 'fruits.' (Programs are, of course, free to also report outcomes on more specialized measures, e.g., 'apples,' 'oranges,' or 'lemons.'). *Well-being* is a commonly advocated outcome for which there are useful measures (cf Drummond et al., 2005). Critics deride well-being as too subjective and as difficult to measure with acceptable reliability and validity.

A cost-inclusive index more agreeable to many than net cost or benefit-cost is *cost per year of life saved*. Diverse programs can be compared in terms of years of life saved. In recognition that the subjective value of a year of life can be less if one is seriously ill and suffering, years of life saved usually are adjusted for *quality*. Accepted algorithms are used to calculate *quality-adjusted life years* (QALYs; cf Gold et al., 1996). Careful research can determine the change (hopefully improvement) in QALY produced by different medical or other interventions (e.g., Freed et al., 2007). Combining measures of gain in QALY with information on the costs of activities that seem to cause those gains results in estimates of the cost of adding a quality-adjusted year of life for patients, or *cost per QALY* (e.g., £/QALY). Similar metrics that include costs can be used to compare the cost-benefit of prevention versus treatment, and health interventions versus efforts to improve quality of life through education, the arts, income support, and more.

A related measure is years of life lost due to disability caused by health and mental health problems and premature death: *DALY*. The DALY combines findings regarding program effects on mortality and morbidity into a single variable: a combination that can facilitate higher-level decisions about funding of public health interventions. Costs of efforts to reduce disability and premature cost often are expressed as *cost per DALY* saved.

Schools of Thought and Leading Advocates

CBA and CEA can be viewed as methods of making difficult decisions in which some individuals gain resources or services at the expense of others. CBA and CEA promise a decision process that is fair because it attempts to weigh social investment against social gain objectively and quantitatively. Critics note that this logical positivist framework has been shown to generate decisions that are fundamentally subjective and that often benefit persons and organizations that have the largest monetary stake in the decision (cf Guba and Lincoln, 1989). The critical, constructivist position makes rejection of cost considerations easier for researchers, service providers, advocates, and consumers who would rather not include costs of services or their monetary consequences decisions about which services to fund and offer, or defund and stop.

Nevertheless, as long as there are limits on the availability, quality, and amount of resources needed to provide services and produce products, the amount of those resources needed for a social program will factor into any realistic decision-making. Evaluations and decisions based on at least the monetary value of those resources will continue to be made, informally at least. CBA and CEA can make the measures and biases of each stakeholder group explicit for all to see. The process by which those resources are valued and monetized can be transparent. In addition, CBA and CEA consider outcomes of human endeavors as well as their costs.

Some critics of human services seize upon program evaluation in general, and CBA and CEA in particular, as a means of challenging the worth of human services and especially entitlement programs. Providing treatment to substance abusers, for example, can be challenged as a waste of public funds. This position has been refuted by findings that substance abuse treatment can cause a subsequent *decrease* in use of expensive health and other human services (Ettner et al., 2005). Concern remains about the cost-benefit and cost-effectiveness of other services, such as free health services and income maintenance. Resolving these issues can be facilitated by cost-inclusive evaluations that are well beyond the scope of this article. Finally, many cost-effectiveness and cost-benefit analysts recognize that their findings should not be the sole criteria on which decisions should be based. Again, we all have fundamental rights which no evaluation should deny any of us.

Changes in Focus or Emphasis over Time

CBA can be viewed as having always been a normal part of most individuals' pragmatic decision-making, in which what is required by an enterprise is compared to what may come of that enterprise. In modern European history, the first formal proposal of CBA was made by Dupuit, a French engineer. While cost-inclusive evaluation in health, substance abuse, and criminal justice has proceeded quickly to approximate the complexity of CBA required to justify dams and other projects by the US Army Corps of Engineers in the 1930s (U.S. Flood Control Act), other human services are only now beginning to measure and compare costs and outcomes carefully and comprehensively. Mental health and substance abuse treatment have a surprisingly long history of advocating, training, and using CEA, and to a lesser degree CBA (e.g., Carter and Newman, 1976; Fishman, 1981; Sorensen and Phipps, 1975; Yates, 1977, 1978). Services for children and families have developed a series of manuals to help programs and researchers assess costs (Dickey et al., 1999; Calculating the Costs of Child Welfare Services Workshop, 2013), but efforts to assess cost-effectiveness and cost-benefits of these and many other human services have only thin histories of cost-inclusive evaluation (cf Levin and McEwan, 2001; Persaud, 2007; e.g., Yates, 2005; Yates et al., 1979).

Some state legislatures in the United States have used CBA to control the amount of funds consumed in health and other human services, especially those that might reduce future costs for criminal justice or health services (cf Pew Charitable Trusts, 2013). Best known are the CBA conducted by Steve Aos and colleagues in response to requests from the Washington State legislature (Washington State Institute for Public Policy, 2011).

As in other areas of program evaluation, 'thumbs up or thumbs down' evaluations have only begun to evolve into 'how can we make it (even) better?' evaluations. Because funding of programs increasingly depends on whether the program meets criteria such as a cost-benefit ratio that is well above 1.0, or a maximum cost per QALY (e.g., Towse, 2009), much expertise has been devoted to standardizing the measurement and calculation procedures. Starting with Gold et al. (1996) and continuing through Yates (1999) and Drummond et al. (2005), consensus is being achieved on more methods. The need for cost-inclusive evaluation has been recognized widely, resulting in an increasing number of submissions to professional journals of CBA and CEA in health and other human services (cf Herman et al., 2009; Yates, 1994).

Current and Emerging Directions in Theory and Research

Utility Assessment

A major advance in cost-inclusive evaluation has been the adoption of outcome measures that are not monetary, but that are standardized across diverse services, such as the QALY and DALY introduced earlier. These measures promise to evolve into a common metric of outcomes that will allow a great diversity of human services to be compared for cost-effectiveness.

Willingness-to-pay to obtain a service and willingness-to-accept withdrawal of a service have been advocated as a measure of the worth of a program since CBA was first formalized (cf Johansson, 1995). Part of the appeal is that this outcome is consumer-driven and could be inexpensive to measure. The willingness-to-pay approach to outcome assessment seems particularly oriented toward business models of social enterprises originating in unabashedly capitalist contexts. While consumers' willingness-to-pay for services and products may be a useful metric when the decision-maker is construed as highly rational and able to pay a range of prices, considerable efforts have been exerted by communications specialists in advertising and politics to distort consumer perceptions.

Willingness-to-pay seems a poor choice for valuing outcomes when consumers seek or are prescribed services precisely because their recent choices do not seem entirely rational, as may be the case for substance abuse, mental health, and criminal justice services. Not having sufficient monetary resources to consider investing in services that may not benefit the consumer for years or decades has led many families to educate only some of their children, a choice that reflects willingness-to-pay but seems a poor choice for society in the long run. Asking economically disempowered consumers to 'imagine' they had a certain amount of resources to invest in services may not generate the rational decisions anticipated, but is likely to uncover deep frustration about economic hardships.

BCA Using Nonmonetary Units

As noted earlier, the primary advantage of CBA over CEA is that costs 'input' and outcomes 'output' are measured in the same units. This allows costs to be subtracted from outcomes to calculate net benefit. It also allows calculation of the *ratio* of outcomes to costs. Costs and benefits of different activities within a program can be assessed in the same nonmonetary units, using similar rating scales for *psychological cost* and *psychological benefit* (cf Yates, 1978). The result provided clinically useful information on client perceptions of which activities were best and worst in a weight-loss program. Similar nonmonetary CBA could be used in group decision-making to allow explication, quantification, and comparison of different program activities consumers and providers perceived to be least difficult and most helpful.

Evidence-Based Practices

Although preferences in training as well as funding have emerged for activities that careful research has shown to produce the outcomes they target, information on the costs of those *evidence-based practices* is rarely reported. The only evidence that seems to matter is nonmonetary effectiveness, and neither benefit nor cost. Many activities, done often and well enough by highly trained professionals working intensively and individually with clients in controlled settings, can be shown to meet criteria for specific outcomes. Calls for including costs and benefits among the evidence for specific practices have gone largely unanswered to date (Yates, 2012a).

For replicability alone, it would seem important to include in descriptions of evidence-based practices careful specification of the types and amounts of each of the key resources needed to implement such a practice. Many human service fields could benefit from reporting the resources needed for a specific evidence-based practice so as to allow researchers and practitioners to estimate the costs of those resources in current monetary values and in their settings (Yates, 1996). Including monetary outcomes, i.e., benefits, when measured, as well as common indices such as QALY and DALY added or saved, could advance the adoption and funding of promising practices.

Requiring inclusion of information on service resources consumed, and their costs, also could encourage researchers to not stop once they have found evidence that a program works. The next step in helping a practice become practical may be to then remove activities that are less responsible for program outcomes, leaving only those that are necessary. This 'efficiency services research,' as it might be called, also could determine the intensity of activities needed to achieve outcomes – for example, whether biweekly rather than weekly meetings were sufficient, and whether trained paraprofessionals rather than psychologists and physician's assistants rather than physicians could implement the essential activities.

Of course, one can also ask for evidence that inclusion of costs in evaluation research is itself effective, and whether it helps to maintain and improve service outcomes. It would be hypocritical not to also ask, is CEA cost-effective? … Is CBA cost-beneficial? These issues of meta-cost-inclusive evaluation research have only begun to be broached empirically (cf Herman et al., 2009; Yates and Taub, 2003).

Methodological Issues and Problems

Costs and Outcomes: Biases in Comprehensive Measurement and Monetization

A persistent and widespread problem in cost-inclusive evaluations is careful measurement of both costs *and* outcomes in the same study and in a standard manner (Siegel et al., 1996). Adding monetary outcomes to reports, and contrasting them with monetary costs, has led to major improvements in use of cost-inclusive evaluation. Most cost-inclusive evaluations and research reports focus on measuring, monetizing, and analyzing either resources invested or outcomes produced but rarely both with equal thoroughness.

Unfortunately, although greater attention has been devoted in the past decade to monetizing outcomes of human services such as substance abuse treatment and prevention of criminal behavior, including careful adjustments of outcomes for quality of research design, costs of such services have been accepted largely at face value. Costs not only need to be adjusted for changes in monetary value over time, such as inflation or deflation, but also can vary between urban and rural regions and can depend on whether the program is offered under the auspices of a larger institution or is a 'stand-alone' operation.

Costs: Monetization of Volunteered and Donated Resources

If costs are undermeasured, programs may seem to be more cost-beneficial, cost-effective, or to have greater cost-utility than is actually the case (Yates, 2012b). Because funding flows more to programs whose outcomes are reported to exceed costs to greater extents, there are strong incentives to measure benefits more comprehensively than costs.

Efforts have been made to standardize measurement of costs using principles of accounting and economics (e.g., Caffray and Chatterji, 2009). French's work is the best known (e.g., French, 2003; McCollister et al., 2009). Paper-and-pencil, computerized, and structured interviews seem to generate more reliable cost data than other methods (cf Zarkin et al., 2004).

Budgets and even accounting records can be poor sources for data on how funds were actually spent, however. Most cost measures used in contemporary CBA, CEA, and CUA neglect entirely what is perhaps the most necessary and widely used resource of all: the time of volunteer providers, and of the clients and family members who transport themselves and others to and from service sites, plus costs associated with transportation vehicles. Accounting records rarely include amounts and estimated values of these 'invisible' but necessary resources.

Some researchers have included measures of volunteered and donated resources (e.g., Yates et al., 2011), and resources devoted to treatment by clients as well (e.g., McCollister et al., 2009). These studies have found substantially different relationships between costs and outcomes when these oft-ignored resources are monetized. For example, cost-inclusive evaluations of consumer-operated services found that the monetary value of volunteered and donated resources exceeded the monetary budget of some consumer-operated human services (Yates et al., 2011). Fears that the cost of these programs would be exaggerated relative to programs that did not include volunteered and donated resources prompted. Yates and colleague Freed to report volunteered and donated resources separately from 'cash costs,' allaying some program's concerns and defusing resistance to collection of resource data.

Costs: Itemization in Reporting to Allow Generalization by Readers

Many articles and manuscripts describe in only a summary manner how cost and benefit data were collected. Reporting of cost evaluation methodology often is so minimal that, if same were done for nonmonetary outcomes, reviewers would reject it immediately. If cost- and benefit-assessment methods were reported in detail sufficient to allow replication by other researchers, cost-inclusive assessments could be more useful to readers and researchers in other countries and at future times. Costs and benefits should be itemized when reported, and quantified with their 'native units,' e.g., hours, square meters, and improvement in days employed, rather than just 'dollars invested' and 'dollars generated and saved.' In addition to reporting the total cost of a treatment program, the specific resources used should also be specified (e.g., X hours per week of direct services by a social worker, for instance) along with the pay rate (salary + benefits) in effect at the time data were collected (cf Sava et al., 2009).

This would allow decision-makers in other monetary systems operating a decade or more after publication of research findings to estimate the costs of each resource used and the value of each benefit generated. Without this, an intervention that has superior cost-benefit or cost-effectiveness to another intervention in one economy may be overgeneralized to be superior in other economies, which have such different pay rates as to reverse which intervention is more cost-beneficial and more cost-effective (e.g., Sava et al., 2009; cf Yates, 2011). Similar itemization of benefits would allow adjustment and potential similar reversals of cost-benefit findings for countries that apply different charges for emergency room, inpatient, and outpatient services.

Outcomes: Monetization of Lifetime Earnings Can Exacerbate Discrimination

Monetization of program outcomes often includes not only savings in future health and criminal justice services that will not be required because of program impact, but also increments in future earnings that result from training, education, or longevity enhancement. This income + benefits calculation is common in evaluations of the worth of a college education as well as estimation of DALY generated by some health programs. If some genders, ethnicities, or age groups use more expensive health or criminal justice services, savings in use of those services would be larger for, and favor focusing efforts on, those particular genders, ethnicities, or age groups.

That may seem fair, and certainly monetizing the expected enhancement of lifetime earnings acknowledges the importance of employment and entrepreneurship. This same approach to monetizing outcomes also could perpetuate or worsen historical discrimination that, for example, pays males more than females, some ethnic groups more than others, and

certain age groups more than others. Increments in lifetime earnings usually favor efforts to enhance income of younger individuals because they have more years of life to accumulate larger totals of improved earnings. Some readers may observe that ageist, sexist, and racist preferences simply reflect the values that society has already placed on the worth of different individuals' lives. Others may react to this issue by rejecting cost-inclusive evaluations entirely. It seems possible, however, that adjustments could be made to wage rates so that the same rate was used for all ages, genders, races. The youth preference inherent in measures of lifetime earnings could be reduced by calculating benefits to include (1) the commonly observed diminution of intervention effects over years of life, (2) savings in services to the aging that may result from their continued employment, and (3) monetary valuation of volunteer activities including child care and mentoring by the aging.

Costs and Outcomes: Statistical Models Capture Uncertainty, but Cost-Inclusive Indices Seldom Do

Even if they include volunteered and donated resources, and benefits assessed from multiple perspectives, cost-inclusive evaluations may report only total costs and total benefits for a program. This highly efficient method does not reflect the dramatic variability of response to program activities between individual clients whether those are persons, families, communities, nations, or regions of the world. Expressing that variability can be problematic, in that it introduces uncertainty and another statistic next to what was a rather neat single number: the benefit-cost or cost-effectiveness ratio. High variability that deviates from the typical bell-shaped curve of the 'normal' distribution, and which does not disappear even with large samples, seems all too typical of health and other human services. This interindividual and intersite variability seldom is reported by researchers or recalled by funders when using cost-inclusive indices to make funding decisions.

Even worse, small differences between mean or median costs, benefits, and net benefit for competing programs are too often interpreted as 'real' because of the dollar or pound signs involved, when the uncertainty in costs and benefits introduces sufficient variability as to render statistical tests of these apparent differences profoundly nonsignificant, or at least the differences in their effect sizes trivial.

Uncertainty of cost-benefit and cost-effectiveness can, of course, be expressed with confidence intervals, and its effects can be examined in analyses that simulate cost and outcome variability. One comparison of special interest that is rarely reported is against a net benefit of 0, and a benefit-cost ratio of 1.00. In the author's experience, some apparently substantial mean and median net benefits and promising benefit-cost ratios prove to be statistically or realistically the same as 0 and 1, respectively!

Benefit-Cost and Cost-Effectiveness Ratios Do Not Describe Relationships Over a Range of Values

A long-recognized but seldom solved problem with ratios that include costs is that they are expressions, which discard information on both costs and outcomes and provide only an expression of the relationship between those, and often just in one specific context. For example, Program A might have cost $100 000 per year and have generated $200 000 in outcomes. It would have a net benefit of $100 000 and a benefit-cost ratio of 2:1. Program B might address the same sort of problem but cost only $50 000 with $150 000 in outcomes. The net benefit is still $100 000 but the benefit-cost ratio is 3:1. Is Program B better than Program A? Maybe. But suppose Program B was now funded for $100 000 rather than $50 000. Could we expect that the same benefit-cost ratio of 3:1 would hold at the higher level of funding? Not necessarily. Typically there are diminishing returns on investments in a given program and a given technology, and even on investments in progressively more impactful technologies (cf Yates, 1978). Ratios of benefit/cost, effectiveness/cost, or cost/QALY are but transient slopes at a particular cost value of a potentially complex and even nonmonotonic cost-outcome function at that one moment of measurement (Yates, 1996).

Maximizing Cost-Benefit, Cost-Effectiveness, and Cost-Utility

Operations research has long been used in military and business settings to maximize outcomes, such as product deliveries, within resource constraints. Operations researchers first construct mathematical models that quantify the contribution to overall outcomes made by each activity within the program. These activity-specific cost-effectiveness or cost-benefit indices are part of this model. The amount of each type of resource used in each activity also is captured in the model, along with the total amount of each resource type that is available in the context in which the program operates. Linear programming then manipulates this resource-activity-process-outcome (RAPO) model to either maximize overall outcome within the resource constraints, or minimize the overall cost of achieving a targeted overall outcome.

Cost-inclusive evaluators have long understood that CEA and CBA can be preliminary stages of operations research (Yates, 1980, 1996). Sufficiently detailed models have only recently been applied to determine the most cost-effective solutions to complex and socially significant health problems such as HIV prevention and treatment (cf Lin et al., 2012). Yates and colleagues (Yates, 1997) have attempted to isolate, measure, and quantify relationships between (1) *resources* consumed by a program, (2) the *activities* that program uses to engage clients so that (3) specific biological, psychological, and social *processes* are instigated, enhanced, diminished, or eliminated such that specific (4) nonmonetary and monetary *outcomes* are achieved, both proximally and distally (cf Yates, 1999). This model promises to facilitate application of linear programming and other operations research solutions to problems of program optimization. It has also facilitated understanding of program failures (Yates, 2002).

Social Return on Investment (SROI): CBA Anew

The SROI approach to improving society is used to fund those health, education, and other efforts that are most likely to yield the best return for society. This form of CBA can be used by private and public agencies at local, national, and international levels to decide how to best allocate their limited resources.

Ideally, SROI analyses identify not only the time to return on investment but also use measures that funders have agreed are valid in a global context. Ideally, SROI allows the same funds to be used repeatedly as monetary resources are converted to societal resources that also have monetary value. The later could include traditional benefit measures such as improved productivity and reduced use of health care, both due to the decreased in physical or mental dysfunction and engendered by enhanced well-being and creativity. Questions can be raised about how soon and how completely these and other social resources can be recognized as monetary resources available for new investments. Arguably, private and public entities should allow and fund all efforts that can be shown to have returns that, after adjustment for any delays, exceed by a confident amount the investment require.

Of course, errors and biases in measurement of benefits and costs noted earlier can substantially distort SROI decisions – a problem that proponents of SROI address by advocating involvement of all stakeholder perspectives, careful monetization, and validation of findings. Eschewing the subjective and measuring only what is 'material' is a principle that risks making decisions that damage health while increasing capital. Fortunately, SROI specialists are both aware of this danger and working on solutions. This is, perhaps, the most novel and promising application of CBA and CEA in contemporary settings.

See also: Alcohol Interventions: Disease Models vs. Harm Reduction; Applied Social Research, History of; Behavioral Economics, History of; Ethical Questions in Social and Behavioral Sciences, History of; Evidence-Based Practice; Implementation Science; Methodology of the History of the Social and Behavioral Sciences; Policy Analysis; Pragmatist Social Thought, History of; Prevention Research; Quantification in the History of the Social Sciences; Science and Politics: Value Neutrality; Social Question: Impact on Social Thought; Utilitarian Social Thought, History of.

Bibliography

American Evaluation Association, 2013. An Evaluation Roadmap for a More Effective Government. http://www.eval.org/d/do/472 (accessed 13.02.14.).

APA Presidential Task Force on Evidence Based Practice, 2006. Evidence-based practice in psychology. American Psychologist 61, 271–285.

Caffray, C.M., Chatterji, P., 2009. Developing an internet-based survey to collect program cost data. Evaluation and Program Planning 32, 62–73.

Calculating the Costs of Child Welfare Services Workshop, 2013. Cost Analysis in Program Evaluation: A Guide for Child Welfare Researchers and Service Providers. Children's Bureau, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.

Carter, D.E., Newman, F.L., 1976. A Client-Oriented System of Mental Health Service Delivery and Program Management: A Workbook and Guide. National Institute of Mental Health, Rockville, MD. DHEW Publication No. (ADM), pp. 76–307.

Dickey, B., Beecham, J.K., Latimer, E., Leff, J.S., 1999. Estimating Per Unit Treatment Costs for Mental Health and Substance Abuse Programs. Human Services Research Institute, Cambridge, MA.

Drummond, M.F., Sculpher, M.J., Torrance, G.W., O'Brien, B.J., Stoddart, G.L., 2005. Methods for Economic Evaluation of Health Care Programmes. Oxford University Press, New York, NY.

Ettner, S.L., Huang, D., Evans, E., Ash, D.R., Hardy, M., Jourabchi, M., Hser, Y., 2005. Benefit-cost in the California Treatment Outcome Project: does substance abuse treatment "pay for itself?" Health Research and Education Trust 192–213. http://dx.doi.org/10.1111/j.1475-6773.2005.00466.x.

Fishman, D.B., 1981. A Cost-Effectiveness Methodology for Community Mental Health Centers: Development and Pilot Test. National Institute of Mental Health, Rockville, MD. DHHS Publication No. (ADM), pp. 81–767.

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., 2003. Drug Abuse Treatment Cost Analysis Program (DATCAP): User's Manual, eighth ed. University of Miami, Miami, FL.

Gold, M.R., Siegel, J.E., Russell, L.B., Weinstein, M.C., 1996. Cost-Effectiveness in Health and Medicine. Oxford University Press, New York, NY.

Guba, E.G., Lincoln, Y.S., 1989. Fourth Generation Evaluation. Sage Publications, Newbury Park.

Herman, P.M., Avery, D.J., Schemp, C.S., Walsh, M.E., 2009. Are cost-inclusive evaluations worth the effort? Evaluation and Program Planning 32, 55–61.

Johansson, P.O., 1995. Evaluating Health Risks. Cambridge University Press, Cambridge, UK.

Levin, H.M., McEwan, P.J., 2001. Cost-Effectiveness Analysis, second ed. Sage Publications, Thousand Oaks, CA.

Lin, F., Terrell, C., Sansom, S., 2012. HIV Prevention Resource Allocation Model: A Mathematical Model Designed for ECHPP Interventions and Lessons from Philadelphia. Presentation at PS12–1201 Grantee Meeting March 26–28, 2012.

Marseille, E., Daandona, L., Saba, J., McConnel, C., Rollins, B., Gaist, P., et al., 2004. Assessing the efficiency of HIV prevention around the world: methods of the prevent AIDS network for cost-effectiveness analysis (PANCEA) project. Health Services Research 39, 1993–2012.

McCollister, K.E., French, M.T., Pyne, J.M., Booth, B., Rapp, R., Carr, C., 2009. The cost of treating addiction from the client's perspective: results from a multimodality application of the client DATCAP. Drug and Alcohol Dependence 104, 241–248.

Persaud, N., 2007. Conceptual and Practical Analysis of Costs and Benefits in Evaluation: Developing a Cost Analysis Tool for Practical Program Evaluation. USA: ProQuest Information and Learning Company.

Pew Charitable Trusts, 2013. States' Use of Cost-benefit Analysis: Improving Results for Taxypayers. July 2013. Downloaded 20.02.14 from: MacArthur Foundation http://pewstates.org/resultsfirst.

Posavac, E.J., 2011. Program Evaluation: Methods and Case Studies, eighth ed. Prentice Hall, Boston, MA.

Rogers, P.J., Stevens, K., Boymal, J., 2009. Qualitative cost-benefit evaluation of complex, emergent programs. Evaluation and Program Planning 32, 83–90.

Sava, F.A., Yates, B.T., Lupu, V., Hatieganu, I., Szentagotai, A., David, D., 2009. Cost-effectiveness and cost-utility of cognitive therapy, rational emotive behavioral therapy, and fluoxetine (Prozac ®) in treating depression: a randomized clinical trial. Journal of Clinical Psychology 65, 36–52.

Shadish, W.R., Cook, T.R., Campbell, D.T., 2001. Experimental and Quasi-Experimental Designs for Generalized Causal Inference, second ed. Houghton Mifflin Harcourt, New York.

Siegel, J.E., Weinstein, M.C., Russell, L.B., Gold, M.R., 1996. Recommendations for reporting cost-effectiveness analyses. Journal of the American Medical Association 276 (16), 1339–1341.

Sorensen, J.E., Phipps, D.W., 1975. Cost-finding and Rate-Setting for Community Mental Health Centers. National Institute of Mental Health, Rockville, MD. DHEW Publication No. (ADM), pp. 76–291.

Towse, A., 2009. Should NICE's threshold range for cost per QALY be raised? Yes. British Medical Journal 339, 181.

United Nations, 2014. Universal Declaration of Human Rights. http://www.un.org/en/documents/udhr/index.shtml (accessed 06.03.14.).

Washington State Institute for Public Policy, 2011. Return on Investment: Evidence-Based Options to Improve Statewide Outcomes. Author, Olympia, WA. Available at: wsipp.wa.gov.

Yates, B.T., 1977. A cost-effectiveness analysis of a residential treatment program for behaviorally disturbed children. In: Mittler, P. (Ed.), Research to Practice in Mental Retardation, vol. 1. University Park, Baltimore, pp. 435–445.

Yates, B.T., 1978. Improving the cost-effectiveness of obesity programs: reducing the cost per pound. International Journal of Obesity 2, 249–266.

Yates, B.T., 1980. Improving Effectiveness and Reducing Costs in Mental Health. Thomas, Springfield, IL.

Yates, B.T., 1994. Toward the incorporation of costs, cost-effectiveness analysis, and cost-benefit analysis into clinical research. Journal of Consulting and Clinical Psychology 62, 729–736.

Yates, B.T., 1996. Analyzing Costs, Procedures, Processes, and Outcomes in Human Services: An Introduction. Sage Publications, Thousand Oaks, CA. Chosen for inclusion in the 2010 edition of Sage Research Methods Online.

Yates, B.T., 1997. Formative evaluation of costs, cost-effectiveness, and cost-benefit: toward cost procedure process outcome analysis. In: Bickman, L., Rog, D. (Eds.), Handbook of Applied Social Research Methods. Sage, Thousand Oaks, CA, pp. 285–314.

Yates, B.T., 1999. Measuring and Improving Cost, Cost-Effectiveness, and Cost-Benefit for Substance Abuse Treatment Programs. National Institute on Drug Abuse, Bethesda, MD. NIH Publication Number 99-4518, 124-page manual, initial press run 10,000 copies. Full copy readable and downloadable from NIDA site: http://archives.drugabuse.gov/IMPCOST/IMPCOSTIndex.html.

Yates, B.T., 2002. Roles for psychological procedures, and psychological processes, in cost-offset research: cost procedure process outcome analysis. In: Cummings, N.A., O'Donohue, W.T., Ferguson, K.E. (Eds.), The Impact of Medical Cost Offset on Practice and Research: Making It Work for You. Context Press, Reno, NV, pp. 91–123.

Yates, B.T., 2005. Cost-effectiveness analysis and cost-benefit analysis. In: DuBois, D., Karcher, M. (Eds.), Handbook for Youth Mentoring. Sage Publications, Thousand Oaks, CA, pp. 525–545.

Yates, B.T., 2009. Cost-inclusive evaluation: a banquet of approaches for including costs, benefits, and cost-effectiveness and cost-benefit analyses in your next evaluation. Evaluation and Program Planning 32, 54–56.

Yates, B.T., 2011. Delivery systems can determine therapy costs, and effectiveness, more than type of therapy. Perspectives on Psychological Science 6, 498–502.

Yates, B.T., 2012a. Quantitative approaches to outcome measurement: a commentary. In: Magnabosco, J., Manderscheid, R. (Eds.), Outcomes Measurement in Human Services, second ed. National Association of Social Workers Press, Washington, DC, pp. 47–57.

Yates, B.T., 2012b. Step arounds for common pitfalls when valuing resources used versus resources produced. In: Julnes, G. (Ed.), Promoting Valuation in the Public Interest: Informing Policies for Judging Value in Evaluation, vol. 133. New Directions in Program Evaluation, pp. 43–52.

Yates, B.T., Haven, W.G., Thoresen, C.E., 1979. Cost-effectiveness analysis at learning house: how much change for how much money? In: Stumphauzer, J.S. (Ed.), Progress in Behavior Therapy with Delinquents. Charles C Thomas, Springfield, IL, pp. 186–222.

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.

Yates, B.T., Taub, J., 2003. Assessing the costs, benefits, cost-effectiveness, and cost-benefit of psychological assessment: we should, we can, and here's how. Psychological Assessment 15, 478–495.

Yeh, S.S., 2009. Shifting the bell curve: the benefits and costs of raising student achievement. Evaluation and Program Planning 32, 74–82.

Zarkin, G.A., Dulap, L.J., Homsi, G., 2004. The substance abuse services cost analysis program (SASCAP): a new method for estimating drug treatment services cost. Evaluation and Program Planning 27, 35–43.

Relevant Websites

http://www.acf.hhs.gov/programs/cb/resource/cost-workgroup – Cost Assessment, Children's Bureau manual for cost assessment, Calculating the Costs of Child Welfare Services., Component-specific costs.

http://www.usda.gov/oce/risk_assessment/ – Government Offices and Agencies US Department of Agriculture.

http://archives.drugabuse.gov/pdf/Costs.pdf – Manuals, NIDA Manual for Measuring Costs, Cost-Effectiveness, Cost-Benefit.

https://research.tufts-nemc.org/cear4/ – Registries, Tufts Cost-Effectiveness Registry.

http://yosemite.epa.gov/ee/epa/eed.nsf/pages/guidelines.html – US Environmental Protection Agency.

https://www.nlm.nih.gov/nichsr/hta101/ta10106.html – US National Institutes of Health Library.