

## **MISUSE: The Shadow Side of Use**

Michael Quinn Patton

*Utilization-Focused Evaluation*

### **Abstract**

Alkin was a pioneer thought leader in emphasizing that evaluators need to be concerned about both use and misuse. This chapter reviews Alkin's contributions to conceptualizing varieties and types of misuse. The article then examines more recently conceptualized approaches, with special emphasis on mechanical use, compliance use, and inappropriate process use. These new directions are placed in the context of the increased politicization of scientific evidence in popular culture, scientific illiteracy generally, antagonism to critical thinking in decision-making, and national political conflict. Concrete examples of misuse of evaluation are incorporated within a framework of examining "the good, the bad, and the ugly" in evaluation misuse and misevaluation.

### **Introduction**

Be wary of the kinds of questions whose findings are particularly amenable to potential misuse.

Marvin C. Alkin (2011, p. 129)

Marv Alkin was a pioneer in studying evaluation use and his contributions shaped the profession and our ongoing attention to use as a priority and standard against which to judge evaluation quality. Less well-known, but no less important, is that he has consistently warned about misuse. That's the focus of my discussion here. Misuse undermines informed

decision-making. To establish context for this discussion, and in celebration of Alkin's important contributions to evaluation, let me take readers on a short trip down memory lane.

Alkin was invited by Sara Miller McCune, co-founder and publisher of Sage Publications, to review the manuscript that became the first edition of *Utilization-Focused Evaluation* (Patton, 1978). He offered a number of helpful suggestions for improving the book, but his most pointed comments were about the importance of paying more attention to misuse. In our initial utilization study on which the book was based, we had not asked questions about misuse. Alkin helped put that issue on the profession's agenda.

In 1990, Alkin edited and published a book on *Debates on Evaluation* based on a three-day symposium where ten "distinguished evaluation professionals gathered together at the UCLA Malibu Conference Facility for...informal discussion on the topic of evaluation utilization" (p. 9). The transcript index shows extensive attention to evaluation misuse (p. 302) and the book includes Alkin's own significant in-depth reflections on the issue, some of which I reproduce here to ensure that it is part of this historical record.

[T]here is substantial literature related to evaluation utilization. In this section of the book, we have examined another related concept -- misuse. As Patton notes, use and misuse are separate dimensions. The use continuum has as its terminus "nonuse." In essence, this continuum defines the extent to which an evaluation is used...Misuse is a different, but related, concept. While nonuse is a measure of degree or magnitude, misuse is a measure of *manner* of use. On one end of this misuse continuum might be found something called "appropriate use." Misuse (or inappropriate use) represents the negatively loaded end of the

continuum. Thus appropriate use/misuse depicts an ethical dimension. (Alkin, 1990, p. 290)

Alkin then introduces the important distinction between misuse and misevaluation. Misevaluation refers to inappropriate acts of an evaluator, for example, “when the technical aspects of the evaluation have not been conducted adequately (e.g., data collection was done poorly or statistical analyses are incorrect)..., when the evaluator fails to understand the evaluation context properly and therefore misdirects the evaluation...[or] when the evaluator fails to recognize properly his or her obligations for appropriate communication to potential users” (Alkin, 1990, p. 290). Misevaluation can lead to misuse. Technically inadequate evaluations may increase the potential for misuse. Failure to focus on the priority information needs of primary intended users may make findings less relevant and therefore less used. Failure to train decision makers in the appropriate ways to use evaluation information can open the door to misuse. The potential for misuse can be reduced when the evaluator takes a proactive stance in identifying potential abuses of the evaluation report and stipulating what the report does not say.

Alkin linked misevaluation and misuse. For example, he offered a blatant example of how misevaluation can lead to concomitant misuse “when the evaluator accedes to decision-maker suggestions and modifies negative or controversial findings. Misuse by the evaluator also takes place, for example, in selective reporting to paint an inappropriate picture of the program. In this instance, the evaluator, having completed an evaluation report, selectively misuses the evaluation information. In many respects, it is difficult to differentiate these kinds of

evaluator misuse from misevaluation because the reporting function, in whole or in part, is a major element of the evaluator's activities” (Alkin, 1990, p. 291).

In synthesizing the Malibu symposium discussions on use and misuse, Alkin was careful to acknowledge contextual and perspective-based complexities. He asked:

- Is nonuse of a poorly done evaluation appropriate nonuse?
- Is nonuse misuse if it is done on purpose?
- If evaluations are not fully used is that misuse? What about partial use?

In considering these questions, it became clear that “one person's misuse is another's sensible administrative practice” (Alkin, 1990, p. 292). To make sense of the complex possibilities, Alkin generated a category system for misuse consisting of four dimensions that aim to distinguish between potential misuse situations: the client's purpose, the quality of the evaluation, users' intentions, and users' technical sophistication. Figure 1 presents this category system.

[Insert Figure 1 about here]

The misuse category system helps distinguish types of misuse. For example, “symbolic use” may actually be misuse when evaluation clients commission an evaluation “purely for symbolic reasons (for political gain, for publicity, to gain funding, to delay action, or to avoid taking responsibility). This kind of symbolic purpose constitutes misuse of evaluation” (Alkin, 1990, p. 292). He elaborated the distinctions with further examples of their application, in so doing illuminating the nuances and complexities of the concepts of misuse and misevaluation.

In instances where evaluation was commissioned with an instrumental purpose in mind (a potential likelihood of use of evaluation information), the

evaluation misuse model is further differentiated by whether the evaluation was done well or poorly. Clearly, nonuse of a poorly done evaluation is highly laudatory, as is use of a well-done evaluation. The other examples require further differentiation. The issue is whether nonuse of well-done evaluations is unintentional or intentional/blatant, the former of which I am willing to acknowledge as simple nonuse, with higher degrees of intentionality classified as mis-use or abuse.

Likewise, informed users should know better than to rely on poorly done evaluations, and such instances I would categorize as misuse (as well as misevaluation). An instance of a use of a poorly done evaluation by an uninformed user I simply call misevaluation. (Alkin, 1990, pp. 292-293)

These reflections, more than two decades ago, have influenced evaluation research, theory, ethics, and practice. Now, fast forward to the present. In his most recent book, *Evaluation Essentials: From A to Z* (Alkin, 2011), he warns students and readers to “**GUARD AGAINST MISUSE.**” He writes:

I have talked about the importance of your actions in helping evaluation use to occur. Of equal importance is your responsibility for guarding against misuse. Misuse occurs when stakeholders modify, misstate, or inappropriately excerpt from the evaluation report. You have an important responsibility for assuring that your evaluation is conveying the information that you intended and not being misapplied in ways not justified by your evaluation. Misuse may start with stakeholders taking your report and simply modifying sections of it. This is

inappropriate. Misuse may occur by stakeholders summarizing elements of the report in ways that are not consistent with what you stated. This is inappropriate.

Misuse may occur by stakeholders when they injudiciously excerpt portions of the report consistent with their beliefs, but not with the tone of the report. This is inappropriate.

And so I ask you to consider use; do all that you can to foster appropriate use. However, be alert to potential misuse. (Alkin, 2011, pp. 211-2).

Alkin's ethical voice comes through strong and clear in his rhythmic repetition of the refrain: *This is inappropriate*. This clarity of voice is needed now more than ever. As attention to issues of accountability, evidence-based decision-making, performance measurement, and evaluation have increased in the deeply divided political landscape of our times, and as the media attend to, highlight, and distort data and findings to feed the beast of the unrelenting 24-hour news cycle, corruption of evaluation becomes an ever-greater threat to the integrity of our profession and the appropriate use of evaluations, both findings and process. In the remainder of this chapter, I'll provide examples of some of those threats.

### **Examples of Evaluation Misuse**

#### **Misuse of Evaluation Process & Findings**

*Evaluation as the Deadly Hatchet*. Katherine Barrett and Richard Greene have been long-time advocates of performance measurement and evaluation through their B&G Report associated with *Governing* magazine. However, they got a glimpse of the shadow side of evaluation during a dinner with a close friend who, they reported, works in a popular government-funded program.

In the past, the measures used to determine success or failure were far from adequate. But then a number of government leaders became eager to de-fund the program. So they instituted a further-reaching system of evaluations, not necessarily to figure out how to make the program work optimally, but to prove that it was ineffective. In the end, the agency wasn't de-funded. But people in the agency learned a negative lesson: Performance evaluations are a "gotcha" exercise, used by people who have a bias. Of course, we don't think this is generally — or even frequently — true. But we do think that every time it happens, it sets the performance-measurement movement back a step.

(Barrett & Greene, 2010, p. 1)

This type of politically-oriented use has been called "legitimative utilization" (Alkin, 2005, p. 435; Leviton, 2003, p. 533) in which evaluation findings are used to support a decision that was actually made before the evaluation was ever conducted or was made without regard to evaluative evidence. This is what the critics of the Iraq War argue happened, namely, that President Bush and his neoconservative advisors had already decided immediately after the 9/11 terrorist attack on the World Trade Center that they would use the attack as justification for invading Iraq and deposing Saddam Hussein. They then set about gathering and presenting selective "evidence" to legitimate that predetermined decision (United States Senate Select Committee on Intelligence, 2004; Hersh, 2003). This happens in a program context when a decision is made to terminate a program and then an evaluation is commissioned for the purpose of legitimating the decision after the fact. Program staff is often fearful of just such an agenda when internal evaluations are commissioned in a time when resources are known to be

constrained and some cuts somewhere will have to be made. To the extent that *legitimative use* is intentionally manipulative and deceptive, it becomes misuse.

***Symbolic Use as Misuse.*** Symbolic use, referenced briefly earlier, refers to token or rhetorical support for an evaluation process or findings but with no real intent to take either the process or findings seriously. Symbolic use has become more prevalent as research and evaluation findings have become increasingly prominent in political dialogue. In the knowledge age, politicians and decision makers have to at least appear to be basing their views on data. This distinction carries a warning to evaluators not to believe naively easily expressed rhetoric about interest in evaluation. Look for evidence of and specific actions in support of evaluation processes and findings; a reasonable evaluation budget and time devoted to the evaluation are prime types of such evidence. Symbolic use constitutes a shrewd political use of evaluation to give the appearance of being an evidence-based decision maker. Symbolic use is not just misuse of evaluation but misuse of scarce public resources. Symbolic use can breed skepticism about the value of evaluation.

***Imposed Use as Potential Misuse.*** The classic three types of use -- instrumental, conceptual, and symbolic -- have long framed inquiries into evaluation use and led to concerns about misuse (Patton, 2005; 2008). Over time, as the field has matured and inquiries into utilization have broadened and deepened, additional distinctions have emerged from research and theory. For example, based on case studies of the use of D.A.R.E. evaluations, Weiss, Murphy-Graham & Birkeland (2005) identified *imposed use* which occurs when those with the power to do so mandate an action based on evaluative judgments; in essence, those at a higher level of authority require a prescribed use by those at a lower level. For example, a federal



agency may require curricula used in schools to be on an approved list of “evidence-based” or evaluated programs in order for school districts to receive funding. Some school districts, they found, reported feeling forced to drop a popular program, despite local support, because it did not qualify as a pre-approved, evidence-based program by the federal authorities.

### **Misuse of Evaluation Process**

*Compliance Use as Misuse.* Compliance use refers to going through the motions to meet an evaluation requirement. The evaluation is required, so it is done, but the motivation is compliance and the implementation is mechanical. A number of colleagues who do evaluations in the federal government have encountered this approach, as have I, especially with regard to mandated PART reviews, a process mandated by the U.S. Office of Management and Budget (OMB) for all federal programs. PART (Program Assessment Rating Tool) was developed to help budget examiners and federal managers measure the effectiveness of government programs. It is a 25-item questionnaire divided into four sections: program purpose and design (5 questions); strategic planning (8 questions); program management (7 questions); and program results/accountability (5 questions). Based on answers to these questions, a score is generated and a program is rated as Effective, Moderately Effective, Adequate, or Results Not Generated. The stakes are high. Results are made public and can affect program budgets and status. So how does mechanical use come into play? A director of a program preparing for a PART says to the evaluator: "Just tell me what I have to do to increase my PART score." Such a director isn't looking to improve the program or make a decision. The object is just to get a decent, acceptable score. The same phenomenon happens in not-for-profit programs when they go mechanically through the motions of complying with a funder's mandated evaluation.

## **Misuse of Evaluation Findings**

*Misrepresentation of Evaluation Findings in Advertising.* Advertising, by its very nature, is unbalanced. Its purpose is to sell a product not to render a balanced empirical report. The maxim *Buyer Beware* certainly applies to interpretation of advertising claims of all kinds. The distortion of advertising claims is not news. What may be news is the extent to which such distortions misuse evaluation findings. New York Times technology reporters Trip Gabriel and Matt Richtel examined the scientific evaluation claims of software and technology companies aimed at educated, and supposedly sophisticated, consumers. What they found were more sophisticated distortions and misrepresentations aimed at people who value data but won't take the effort to look at original sources. The headline for their investigative report was: *Inflating the Software Report Card* (Gabriel & Richtel, 2011).

### **Factors Contributing to Misuse**

#### **Corruption of Performance Measures**

One common corruption sequence works like this: What gets measured gets done. Measure the wrong thing, or a corollary, use simplistic measures for complex phenomena. At the same time, set punitive consequences for missing simple performance targets. The result will be misuse of evaluation, everything from fudging and faking numbers to meet targets to outright cheating, all of which has been in evidence in the No Child Left Behind scandals.

A particular frightening and egregious example was recently exposed in the New York City narcotics squad. Steve Anderson, a former undercover police officer, testified how rules were trimmed, broken or ignored so that narcotics officers could make their monthly quotas of arrests or buys. He testified that drugs obtained during arrests were held back, what were called

“spare drugs,” to “plant on people when a narcotics officer needed a productivity boost.” It was called "attaching bodies" to the drugs. He testified that during some four years of undercover work he had “become numb to the corruption.” He continued: "It was something I was seeing a lot of, whether it was from supervisors or undercovers and even investigators" (Dwyer, 2011, p. A20). This was reported in the press as a police corruption scandal, but it is also a story of misuse of evaluation performance indicators.

### **Mechanistic Use**

Closely related to imposed use is *mechanistic use*. This occurs when a major decision follows attainment of some rigid fixed performance target. Mechanistic use aims to remove data interpretation and evaluative judgment in order to go directly from finding to action. Perhaps the best-documented example is the sentencing guideline of *3-strikes-and-you're-out*. This refers to the judicial sentencing mandate that three felony convictions will automatically mean life in prison. A review by the U.S. Department of Justice in 2000 found that 24 states and Congress have passed such legislation. The Executive Summary states:

The report proffers that this form of legislation was carefully crafted to be largely symbolic. However, the gross errors in predicting the impact of these and other laws by some of the most prestigious researchers underscore how little we know about change within the criminal justice system. (Austin, Clark, Hardyman & Henry, 2000, Executive Summary).

In 1980 in *Rummel v. Estelle*, the Supreme Court upheld a life sentence for a third-strike fraud felony in Texas, which arose from a refusal to repay \$120.75 paid for air conditioning repair that was subsequently considered unsatisfactory. Thirty years *New York Times* columnist

David Brooks (2011) has opined that “sentencing guidelines are out of control.” He cited as examples people who are sentenced to life without parole for offenses committed while they were kids. He continued: “Similarly, there was the case of the young man who was sent away for life for having images of child pornography on his computer. I’m as revolted as the next person, but the penalty for looking at images should not be greater than the penalty for murder. We should be putting more power in human discretion and less in rigid codes.”

The performance targets at which schools were put on probation under No Child Left Behind administrative rules were similarly implemented without regard to context and without deliberation. Implementation was mechanistic. This was done in the name of fairness, that is, creating a rigid uniform standard that applied equally to all. The actual result of this performance target policy, as demonstrated in the widespread cheating scandals, is that the probation criteria were perceived as unfair precisely because they failed to take into account a school’s context, community situation, efforts at reform, demographic composition, resources, and other factors that could, and should, affect an evaluative judgment. Mechanistic use can be misuse.

### **Single Narrative Simplicity**

Eleanor Chelimsky, former AEA president and retired U.S. Assistant Comptroller General for Program Evaluation has identified the single narrative as a form of evaluation misuse that she believes is on the rise.

What do I mean by the single narrative? Well, it’s an artificially simple idea, or cause-and-effect relationship, neither of which is established by considering all the available evidence, but rather by suppressing that part of the evidence, which counters some pre-determined agenda. So the single narrative is

a false distillation from complexity; its agenda is typically about politics or profit-making; and it often involves lying by omission, saying we don't know something when we do, or that we do know something when we don't. And usually, it's also about taking action – in the form of a government policy or program – that embodies this carefully customized reality...

What I do want to address here is the deliberate falsehood, the disingenuous political spin that affects evaluation both at the start of a study, by rendering all our arduous methodological planning irrelevant, and at the study's end, with the refusal of sponsors and users to listen to any evaluation results other than those which fit some already-established, inflexible position. (Chelimsky, 2011, p. 1).

Chelimsky, as one of evaluation's most distinguished pioneers, has been an astute observer of evaluation patterns in the federal government for 50 years. It merits our full attention and concern, then, when she asserts: "The single narrative seems to be growing in momentum, in application, and in use." (p. 2). She also asserts that the effects on evaluation as it spreads are likely to be substantial. She attributes the spread of the single narrative to several interrelated factors: (1) political ideologies at the extremes dominating policy debates; (2) the increase in Congressional partisanship to the point where strong scientific evidence is routinely rebuffed; (3) a general failure to attend to data to such an extent that even "the best data seem to have lost their ability to persuade;" and (4) the Internet "fueling a trend in which everyone feels free to invent his own facts" (Chelimsky, 2011, p. 4). The corruption of the intelligence community in orchestrating allegations of weapons of mass destruction in Iraq to justify the invasion there is

but one prominent example she cites. Another example is the single narrative that bad teachers are the single cause for poor student achievement and the accompanying narrative that inadequate teachers are protected by powerful teachers' unions. Countering the simple-mindedness of the single narrative requires dealing with complexity, including a much-needed capacity among evaluators to apply complexity concepts to enhance appropriate evaluation use and deal with complex narratives (Patton, 2011). The trend away from engaging complexity and toward single narrative simplicity may well be undergirded by larger cultural and societal forces that oppose reasoned dialogue, to wit, *the evangelical rejection of reason*.

### **The Evangelical Rejection of Reason**

Karl W. Giberson is a former professor of physics, and Randall J. Stephens is an associate professor of history, both at Eastern Nazarene College. They wrote, as evangelicals, an opinion piece, published in *The New York Times*, on “the evangelical rejection of reason.”

The Republican presidential field has become a showcase of evangelical anti-intellectualism. Herman Cain, Rick Perry and Michele Bachmann deny that climate change is real and caused by humans. Mr. Perry and Mrs. Bachmann dismiss evolution as an unproven theory...

The rejection of science seems to be part of a politically monolithic red-state fundamentalism, textbook evidence of an unyielding ignorance on the part of the religious. As one fundamentalist slogan puts it, “The Bible says it, I believe it, that settles it.” ...But when the faith of so many Americans becomes an occasion to embrace discredited, ridiculous and even dangerous ideas, we must not be afraid to speak out... (Giberson & Stephens, 2011, p. A27).

This rejection of science and critical thinking has been labeled *denialism*. This term is being used among science journalists to describe public reactions to politically and socially charged scientific issues like climate change, evolution, vaccines, AIDS, the 9/11 attacks, and even the Holocaust. The lack of what we, as evaluators, would call evaluative thinking was the theme of the World Conference of Science Journalists in Doha, Qatar, June 27-29, on "Journalism in the Age of Denial." One panel, organized and moderated by Cristine Russell, president of the Council for the Advancement of Science Writing, former *Washington Post* science reporter, and senior fellow at Harvard's Belfer Center for Science and International Affairs, began by positing that like scientists and educators [and, dare we say, evaluators], science journalists operate under the basic assumption that more and better information is desirable to inform public dialogue. Science journalists aim to accurately communicate scientific information to the public to enhance understanding and support democratic decision-making. But the panel of science journalists concluded that denialists don't share these values or this perspective. Consider the conclusions of science journalist Shankar Vedantam, author of *The Hidden Brain* who recently moved from the Washington Post to National Public Radio. He asserted: "Telling people with strong partisan beliefs that something is not true rarely convinces them. In fact, refutation causes the belief to go up" (as quoted by Frazier, 2011, p. 5.)

In essence, providing people who hold strong beliefs reliable information that might undermine those beliefs causes them to cling to those beliefs even more strongly. As evaluators we've long seen this phenomenon with the denial of findings that D.A.R.E. doesn't work to reduce drug use among middle school children (Patton, 2012, chapter 1). A summary of the experiences and views of other science journalism panelists illuminates the environment in

which evaluators and now working and the basis for increased resistance to evaluation findings specifically and evaluative thinking more generally.

Debra MacKenzie, a writer in the Brussels office of *New Scientist* and author of a 2010 cover article "Living in Denial," reinforced that point. "Most denialists simply think the way most people think -- in terms of feelings, familiar stories, and their own group identity," she emphasized. When they hear a conflicting claim, cognitive dissonance takes over. "It is far easier to deny the science than to accept that your whole worldview is wrong." Yves Sciamia, a freelancer from France, wondered whether the honeymoon between science and society earlier in the twentieth century is now coming to an end. "Scientist-bashing is becoming more common," Sciamia noted, whereas it was almost unheard of decades earlier.

The main concern of Philip Hilts, head of the Knight Science Journalism Fellowships at MIT, is the "deliberate misuse of nonsense." He said tactics used by the tobacco companies from 1953 to 1994 to try to counter the evidence that smoking causes cancer are now being used by oil companies to try to counter the evidence of climate change. The tobacco companies tried to assert that there was always doubt. If necessary, they would "buy a scientist to say so." Now, with climate change, Hilts said, the vested interests are again "deliberately manipulating for gain." He said the role of science journalists is to "chase down the folks" who are misusing and misrepresenting the scientific evidence.



Panelists seemed to agree that denialism is worsening. "It is hard to believe that denialism is not expanding," Vedantam said. He noted that in regard to global climate change, disbelief in the United States is going up. (Frazier, 2011, pp. 5-6).

### **Misuse on the Rise**

In *Utilization-Focused Evaluation* (Patton, 2008), I proffered the following proposition: *As use increases, misuse will also increase*. When people ignore evaluations, they ignore their potential uses as well as abuses. As evaluators successfully focus greater attention on evaluation data and increase actual use, there may be a corresponding increase in abuse, often within the same evaluation experience. Evaluation pioneer and visionary Donald T. Campbell formulated a discouraging law along these lines that has come to be known as Campbell's Law:

The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor (Campbell, 1988, p. 306).

Sources of misuse include hard-core politics, asking the wrong questions, pressures on internal evaluators to present only positive findings, petty self-interest, and ideology. Misuse, like use, is ultimately situational. Consider, for example, the case of an administrator who blatantly squashes several negative evaluation reports to prevent the results from reaching the general public. On the surface, such an action appears to be a prime case of misuse. Now consider the same action (i.e., suppressing negative findings) in a situation where the reports were invalid due to poor data collection. Thus, misuse in one situation may be conceived of as

appropriate nonuse in another. Intentional nonuse of poorly conducted studies can be viewed as appropriate and responsible.

Misuse can be either intentional or unintentional. Unintentional misuse can be corrected through the processes aimed at increasing appropriate and proper use. Intentional misuse is an entirely different matter that invites active intervention to correct whatever has been abused, either the evaluation process or findings. As with most problems, correcting misuse is more expensive and time-consuming than preventing it in the first place.

Working with multiple users who understand and value an evaluation is one of the best preventatives against misuse. Allies in use are allies against misuse. Indeed, misuse can be mitigated by working to have intended users take so much ownership of the evaluation that they become the champions of appropriate use, the guardians against misuse, and the defenders of the evaluation's credibility when misuse occurs.

Policing misuse is sometimes beyond the evaluator's control, but to the extent possible and realistic, professional evaluators have a responsibility to monitor, expose, and prevent misuse (Patton, 2005).

### **Nonuse: Appropriate versus Inappropriate Occurrences**

The utility standards of the profession make it clear that a good evaluation is one that is used. This volume connects evaluation use and informed decision-making. Some use is good; more use is better. Appropriate and intended use by intended users is best. Misuse is bad. And nonuse?

From a utilization-focused evaluation perspective, nonuse represents some kind of failure in the evaluation process. We often lay that failure at the feet of resistant or unappreciative

stakeholders, but it can also be the evaluator's fault. As discussed earlier, *nonuse due to misevaluation* (Alkin, 1990, pp. 290-294; Patton 2005, p. 254) can be viewed as appropriate nonuse because of weak evidence, a late report, poor evaluator performance, or other failures of the evaluator to adhere to the profession's standards and principles. This is “justified nonuse” (Cousins & Shulha, 2006, p. 282). In contrast, *political nonuse* occurs when the findings are ignored because they conflict with a potential user's values, prejudices, preferences, and predisposition -- so the evaluation is just simply ignored. Utilization-focused evaluation attempts to reduce political nonuse by creating a climate and process in which those involved are willing and prepared to examine their basic assumptions and incorporate evidence into their understandings, even when they had hoped for, or would have preferred, different results.

*Aggressive nonuse*, or calculated resistance, refers to situations where an evaluation or evaluator is attacked and use is undermined because the results conflict with or raise questions about a preferred position. Resistance to evaluation findings can be a specific example of the more general phenomenon of resistance to change. A major reason for identifying and involving primary intended users in the evaluation is to anticipate and short-circuit inappropriate and specious attacks, or at least to have allies among informed and credible intended users in fending off such politically motivated attacks.

Most resistance to evaluations is behind-the-scenes, but occasionally political reports grab media attention and the whole world gets to watch the circus of attacks and counter-attacks. A prominent example was the May, 2005 release of a report by the human rights organization Amnesty International on conditions in the U.S. military prison at Guantanamo Bay in Cuba where alleged terrorists were being held. The report, citing interviews with prisoners and people

who had been inside the prison, concluded that prisoners had been mistreated and called for the prison to be shut down. The report got considerable international media attention. Amnesty International has an explicit agenda and its recommendation to close the Guantanamo facility could be expected, but the cases cited and interview results were viewed as credible by some reporters, so the Bush Administration needed to make a response. The tone of the response gives a flavor of the rhetoric that can accompany an aggressive attack on disputed and unwelcome evaluation conclusions. President Bush, addressing a news conference at the White House on May 31, 2005, said the Amnesty document was an “absurd report. It’s absurd. It’s an absurd allegation. The United States is a country that promotes freedom around the world.” He went on to attack the investigation's methods and resulting data asserting that the Amnesty allegations were based on interviews with detainees who hated America and were trained to lie. President Bush's remarks were echoed by Vice President Dick Cheney, who said that same day in a videotaped interview with CNN's Larry King, "Frankly, I was offended by it. For Amnesty International to suggest that somehow the United States is a violator of human rights, I frankly just don't take them seriously.”

In the early 1970s I was involved in an independent survey of teachers in Kalamazoo, Michigan with funds from the local and national education associations. The School District refused to cooperate with the study and when the results came in showing very low morale, widespread complaints about working conditions, a dysfunctional accountability system, and allegations of administrative abuses, the Superintendent publicly attacked the findings, calling them "absurd." He attacked my integrity, saying I was an out-of-state paid-gun-for-hire, and further asserted that the teachers association instructed teachers how to respond. He dismissed

the results out of hand. Fortunately, the school board members actually read the report, including pages of in-depth quotations from teachers and documented cases of problems. The school board made instrumental use of the report by requiring major administrative changes in the District and, subsequently, the superintendent "resigned." (For details, see Patton 2002, pp. 17-20.)

The point: Evaluation is a political activity and as the varieties of use, nonuse, and misuse illustrate, utilization is also a political activity -- and sometimes the politics gets rough. This work is not for the feint of heart; it's not just an academic exercise. The stakes can get very high, very fast.

### **Overuse**

Let me close with an emergent and unexpected phenomenon I've come to call *overuse*. Having spent my evaluation career promoting evaluation use, it came as a surprise to suddenly find that overuse was becoming a problem. *Overuse* occurs when too much emphasis is placed on evaluation findings. For example, weak evaluation results are overused when treated as if they are definitive, or imposed use becomes overuse when there is insufficient evidence to generalize findings and justify the top-down mandate for compliance, or there is lack of attention to local conditions. This latter overuse can occur when supposed "best practices" are universally mandated (Patton, 2001). Mechanistic use, described earlier, can be a form of overuse. Concern about overuse is ironic since the profession has been dominated by concern about under-use and nonuse. But, as in much of life, you can have too much of a good thing. An unintended consequence of all the focus on increasing use may have contributed to overuse and misuse. We must guard against both.

## References

- Alkin, M.C. (1990). *Debates on Evaluation*. Newbury Park, CA: Sage.
- Alkin, M.C. (2005). Utilization of evaluation. In S. Mathison (Ed.), *Encyclopedia of Evaluation*. pp. 434-436. Thousand Oaks, CA: Sage.
- Alkin, M.C. (2011). *Evaluation Essentials: From A to Z*. New York, NY: Guilford.
- Austin, J., Clark, J., Hardyman, P. & Henry, D.A. (2000). "Three strikes and you're out." The implementation and impact of strike laws. Washington, DC: The National Institute of Justice, Office of Justice Programs, U.S. Department of Justice.
- <https://www.ncjrs.gov/pdffiles1/nij/grants/181297.pdf>
- Barrett, K. & Greene, R. (2010). Performance reporting. *B&G Report, Governing*, September 9, p. 1.
- <http://www.governing.com/columns/bg-report/hospital-errors-public-private-partnerships-flexibility-vacation-time-performance-evaluations-cell-phones-government-management-news.html#item3>
- Brooks, D. & Collins, G. (2011). The horror show at Penn State. *New York Times*, November 9.
- <http://opinionator.blogs.nytimes.com/2011/11/09/the-horror-show-at-penn-state/?hp>
- Campbell, D.T. (1988). *Methodology and Epistemology for Social Science: Selected Papers*. E.S. Overman (Ed.). Chicago, IL: University of Chicago Press.
- Chelmsky, E. (2011, November). Evaluation and the single narrative. Presentation at the American Evaluation Association Conference, Anaheim, California.
- Cousins, J.B. & Shulha, L.M. (2006). A comparative analysis of evaluation utilization and its

- cognate fields of inquiry: Current issues and trends. In I.F. Shaw, J.C. Greene & M.M. Mark (Eds.), *The Sage Handbook of Evaluation: Policies, Programs and Practices*. pp. 266-291. Thousand Oaks, CA: Sage.
- Dwyer, J. (2011). The drugs? They came from the police. *New York Times*, October 14, p. 20A.
- Frazier, K. (2011). The age of denialism: When beliefs trump scientific facts. *Skeptical Inquirer*, 35(6): 5-6.
- Gabriel, T. & Richtel, M. (2011). Inflating the software report card. *New York Times*, October 9, p. A1.  
<http://www.nytimes.com/2011/10/09/technology/a-classroom-software-boom-but-mixed-results-despite-the-hype.html>
- Giberson, K.W. & Stephens, R.J. (2011). The evangelical rejection of reason. *New York Times*. October 18, 2011, p. A27.  
<http://www.nytimes.com/2011/10/18/opinion/the-evangelical-rejection-of-reason.html?hp>
- Hersh, S.M. (2003). The stovepipe: How conflicts between the Bush Administration and the intelligence community marred the reporting on Iraq's weapons. *The New Yorker*, October 27, [http://www.newyorker.com/fact/content/?031027fa\\_fact](http://www.newyorker.com/fact/content/?031027fa_fact)
- Leviton, L. (2003). Evaluation use: Advances, challenges and applications. *American Journal of Evaluation*, 24(4): 525-535.
- Patton, M.Q. (1978). *Utilization-Focused Evaluation* (1<sup>st</sup> Ed.). Newbury Park, CA: Sage.
- Patton, M.Q. (2001). Evaluation, knowledge management, best practices, and high-quality lessons learned. *American Journal of Evaluation*, 22(3): 329-336.
- Patton, M.Q. (2002). *Qualitative Research and Evaluation Methods* (3rd Ed.). Thousand Oaks,

CA: Sage.

Patton, M.Q. (2005). Misuse of evaluations. In S. Mathison (Ed.), *Encyclopedia of Evaluation*. pp. 254-5. Thousand Oaks, CA: Sage.

Patton, M.Q. (2008). *Utilization-Focused Evaluation* (4<sup>th</sup> Ed.). Thousand Oaks, CA: Sage.

Patton, M.Q. (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York, NY: Guilford.

Patton, M.Q. (2012). *Essentials of Utilization-Focused Evaluation*. Thousand Oaks, CA: Sage.

United States Senate Select Committee on Intelligence. (2004). *Report of the Select Committee on Intelligence on the U.S. Intelligence Community's Prewar Intelligence Assessments on Iraq*. Washington, D.C.: United States Senate.

Weiss, C.H., Murphy-Graham, E. & Birkeland, S. (2005). An alternate route to policy influence: How evaluations affect DARE. *American Journal of Evaluation*, 26(1): 12-30.



## Figures

Figure 1: Alkin's category system for misuse (Alkin, 1990, p. 293)

