

Citizens and Evaluation: A Review of Evaluation Models

Pirmin Bundi¹  and Valérie Pattyn²

American Journal of Evaluation
1-30

© The Author(s) 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/10982140211047219

journals.sagepub.com/home/aje



Abstract

Evaluations are considered of key importance for a well-functioning democracy. Against this background, it is vital to assess whether and how evaluation models approach the role of citizens. This paper is the first in presenting a review of citizen involvement in the main evaluation models which are commonly distinguished in the field. We present the results of both a document analysis and an international survey with experts who had a prominent role in developing the models. This overview has not only a theoretical relevance, but can also be helpful for evaluation practitioners or scholars looking for opportunities for citizen involvement. The paper contributes to the evaluation literature in the first place, but also aims to fine-tune available insights on the relationship between evidence informed policy making and citizens.

Keywords

evaluation theory, models, citizens, review

Stop! First, let me make something clear; evaluators do not deal with stakeholder groups. Evaluators work with individuals who may represent or be part of different stakeholder groups. I think it is best to work with individuals—specific individuals (Alkin & Vo, 2017, p. 51).

Introduction

Whereas the importance of knowledge has been recognized for centuries (cfr. for instance Francis' Bacon (1605) famous aphorism: knowledge is power), its societal role has changed dramatically in recent decades. The preponderant source of wealth is no longer merely industrial and product related, but is knowledge related. To be able to compete and succeed in this globalized world, societies are increasingly dependent on the knowledge of stakeholders to drive innovations and entrepreneurship. Thus, previous research has discussed how individuals and organizations can participate in evaluation processes (Brandon & Fukunaga, 2014; Cousins, 2003; Cousins & Whitmore, 1998; Greene, 1988; Sturges & Howley, 2017; Taut, 2008; Whitmore, 1998). Cousins and Earl (1992,

¹ Swiss Graduate School of Public Administration, University of Lausanne, Lausanne, Switzerland

² Institute of Public Administration, Leiden University, The Hague, The Netherlands

Corresponding Author:

Pirmin Bundi, Swiss Graduate School of Public Administration, University of Lausanne, Rue de la Mouline 28, CH-1015 Lausanne, Switzerland.

Email: pirmin.bundi@unil.ch

p. 397) argue that the participation of stakeholders is an elementary element in order to understand evaluation utilization.

Starting in the 1980s, participatory methods became more popular in evaluations by “taking more inclusive, rights-based approaches to the design, implementation, monitoring and evaluation of community-based development interventions” (Kibukho, 2021, p. 3). In this context, participation became an instrument that promised to transfer power to the less privileged, and provided them the opportunity to engage in the evaluation process (Cousins & Chouinard, 2012; Hilhorst & Guijt, 2006). While literature on stakeholder participation has a long tradition in research on evaluation, only few studies have discussed the participation of individual citizens in particular. Being often the main beneficiary group of social interventions, citizens de facto have a stake in many interventions and their evaluations, even as “ordinary” or “lay” or “unaffiliated” individuals.

As a matter of fact, in evaluation scholarship, the term “citizen” is often named in one single breath with “stakeholders,” which makes it difficult to delineate both concepts from each other. The often inconsistent use of both terms, and the lack of explicit conceptualizations, contributes to this challenge. With Alkin and Vo (2017, p. 51), we adopt a rather general definition of *stakeholders*, and define them as all individuals who have an interest in the program that is evaluated. This includes clients of the evaluation, program staff and participants, and other organizations. Hanberger’s (2001) distinction between active and passive stakeholders is most insightful in this regard. While active stakeholders, or key actors, will try to influence a policy or program at different stages, passive stakeholders are affected by the policy or program, but do themselves not actively participate in the process. In his understanding, the evaluator needs to recognize, and deliberately include the interest of passive stakeholders, otherwise the effects and value of the policy for inactive or silent stakeholders will be overlooked (Hanberger, 2001, p. 51).

In line with Alkin and Vo’s definition, and consistent with Plottu and Plottu (2009) and Falanga and Ferrão (2021), it can be argued that citizens are a subgroup of stakeholders. Citizens are typically conceived as functional members of a society by virtue of living within it and being affected by it (Kahane et al., 2013, p. 7). Different than organized stakeholder groups such as interest organizations, professional groups or public and private organizations that seek to promote the interest of a limited group, we consider *citizens* as individual persons without an institutional or public mandate, and which are not member of an organized group (such as a political party). Whether citizens are active or passive stakeholders can be assumed to depend on the evaluation at stake, and of the opportunities provided by the evaluator and not least the evaluation model.

Theoretically, the added value of individual citizen participation for evaluation can be justified by referring to three key democratic values: legitimacy, effective governance, and social justice (Fung, 2015). First, citizen participation holds the promise to enhance legitimacy. Citizens may advance interests that are widely shared by other citizens (Bäckstrand, 2006; Fung, 2015). Essential about lay or local knowledge is that it is embedded in a specific cultural and often also practical context (Juntti et al., 2009, p. 209). The input of citizens in the evaluation of interventions can therefore point at communal values that are broadly shared among local communities (Schmidt, 2013), and of which experts and evaluators themselves may not always be aware. Related to this, involving citizens can have epistemological advantages: citizens can be more open to new inputs, and are more aware to how social interventions work in particular communities (Fischer, 2002; Fung, 2015). A belief in communal and local knowledge is also one of the reasons explaining the increased investments in citizen science in recent decades (Irwin, 1995). Secondly, citizen participation can also foster effective interventions, particularly when so-called multisectoral wicked problems are at stake. Citizens, other than political actors for instance, may be well placed to assess trade-offs between ethical or material values; or may frame a policy problem in a more viable way than experts (Fung, 2015). Citizens can advance new viewpoints or an alternative perspective on trade-offs between different types of values, which can foster the validity of certain policies (Juntti

et al., 2009). Finally, citizen participation has the potential to mitigate social injustice (Fung, 2015). From a social justice lens, citizens can bring certain undemocratic biases to the surface in an evaluation.

Despite this potential, little is known as to how much room evaluation theorists have given to citizens in their models. As evaluation theorists differ in their context and evaluation practice, their views on the purpose and the role of citizens also differ fundamentally. Which evaluation theories indeed consider citizen involvement? How and at which stage of the evaluation process are citizens having a role? To date, the evaluation literature does not include such a systematic assessment. As a consequence, this can make it difficult for evaluation commissioners and decision makers to know which models are appropriate for citizen involvement, should they be willing to engage in this. This article addresses this gap by presenting a review of citizen involvement in the main evaluation models circulating in the field. We distinguish between effectiveness models, economic models, and actor-centered models, and studied the state of citizen involvement per stage of the evaluation process. As such, we account for a comprehensive and nuanced assessment on the role of citizens. Method-wise, we triangulated an analysis of the original sources in which the models are outlined, with an expert survey. Our aim is not to make a normative or judgmental assessment of the evaluation models, but to provide a toolbox which can be of use for evaluation practitioners and scholars looking for opportunities for citizen involvement. The article also fine-tunes available insights on the relationship between evaluation and citizens more in general.

The article is structured as follows: the next section sets the stage to introduce the typology of evaluation models, which we use as a heuristic to analyze the role of citizen involvement. The results of the document analysis, and expert survey are presented in section four. The last section summarizes our findings and discusses the implications for evaluation practice.

Evaluation Models

Evaluation models—sometimes referred to as theories or approaches—typically prescribe specific steps that an evaluator is expected to follow towards a particular goal that has been specified in the beginning of the evaluation (Alkin, 2017, p. 141). Of course, there is no such thing as “the” evaluation model. This is why previous scholars have tried to collect and classify different evaluation models. According to Madaus et al. (2000, p. 19), evaluation models are not directly empirically verifiable for a given theory. They are rather to be understood as an evaluation scholar’s attempt to characterize central concepts and ideal typical procedures which can serve as guidelines in evaluation practice. In general, the aim of a classification or a taxonomy is to better understand core principles of different evaluation theories. These especially become clear when defining principles and characteristics are contrasted (Contandriopoulos & Brousselle, 2012, p. 67). Just as there exist many evaluation models, there are quite a number of evaluation taxonomies, all prioritizing different dimensions. For the purposes of our article, we rely on the taxonomy by Widmer and De Rocchi (2012), which is itself based on taxonomies of Vedung (1997) and Hansen (2005).

Vedung (1997, 2004) distinguishes between three major models that constitute the backbone for several subsequent taxonomies, including the one that we adopted for our analysis: (1) substance-only models that primarily address the substantive intervention content, outputs and outcomes, (2) economic models that focus on costs, and (3) procedural models that put intermediary values such as legality, representativeness, and participation in the focus of the evaluation. Hansen (2005) has further built on this overview, and provides a more systematic and fine-grained taxonomy. In comparison to Vedung, she presents a typology of six evaluation models: results models, process models, system models, economic models, actor models, and program theory models. Widmer and De Rocchi (2012), finally, have made an attempt to integrate these typologies, and present a taxonomy with three basic types of models (see Figure 1). In doing so, they make a distinction between

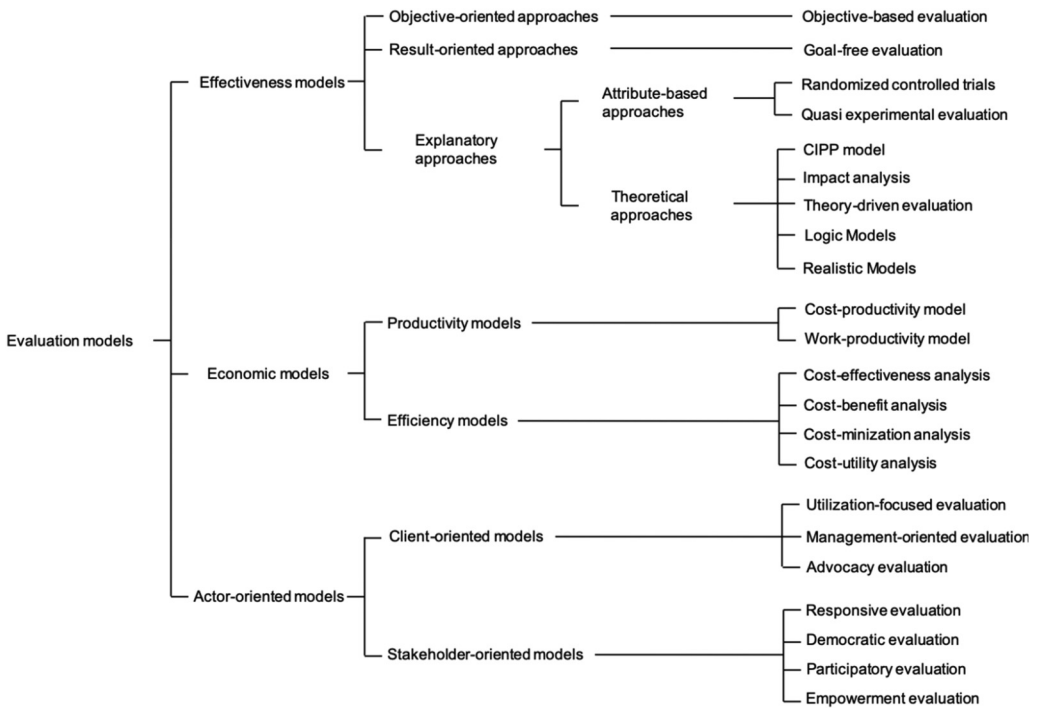


Figure 1. Taxonomy of Evaluation Models By Widmer and De Rocchi (2012, p. 52).

models that focus on the impact of a program (effectiveness models), the efficiency (economic models) or the interests and needs of actors involved and affected (actor-oriented models). While effectiveness models focus on the effects of a program and address only the results of alternative interventions, economic models account for the relationship between program effects and the costs. Finally, actor-oriented models put the emphasis on the actors involved, and were introduced as a separate category to account for the most recent developments in research about evaluation. Figure 1 presents their detailed taxonomy, with 22 models being situated in one of the three overarching types.

Naturally, the position of specific models in the typology can be subjected to debate. For instance, depending on one's approach to advocacy evaluation (compare Sonnichsen's seminal notion (1989) with the work of Julia Coffman (2007)), this model can be situated in a different category. Contandriopoulos & Brousselle (2012: 68) argue that models are "intellectually slippery beasts" which make it challenging to position them in certain categories. Despite these constraints, which are inherent to any typology, the typology by Widmer and De Rocchi (2012) presents a useful heuristic tool for the purpose of our study. Our ambition is not to position these evaluation models definitively, but rather to have a help tool for classifying different evaluation models and their considerations of citizens' interest. First, to the best of our knowledge, Widmer and De Rocchi's taxonomy attempts to be all-inclusive and does not focus on individual policy domains or disciplines. This makes it suitable for all kind of social interventions. Second, the taxonomy makes a more fined-grained distinction between different actor-oriented models, which are in particular important for the participation of citizens. Third, this taxonomy puts predominant emphasis on the original models, and not on later variations or interpretations. While we acknowledge that models can and have been adapted over time, our aim is to analyze the role of citizens as it has been conceived in the "core

idea” of the models. This core idea, as we believe, is best reflected by the original models, and not by their later developments or further developments. Evaluation models are supposed to guide practice. As a consequence, they have been naturally adapted over time by evaluators, to make them fit for a particular local purpose or context. With such empirical applications sometimes very much departing from the original “pure” nature of the models, and being very diverse in nature, we see most added value in staying close to the original purpose of the models, which best represent the common denominator inspiring later empirical applications that originated from this same trunk.

The key question is then how “citizen involvement” is understood in each of these models? To be clear, and in contrast to the concept of citizen science (Irwin, 1995), we do not consider citizens as evaluators themselves per se, but rather review which role is given to citizens in evaluations in general. We simply review any role given to citizens to which reference is made in the models. To account for a comprehensive assessment, we do not restrict our analysis to a particular aspect of the evaluation process, and consider all stages of a typical policy evaluation. With Vedung (1997) and Hansen (2005) we break the evaluation process down in five basic stages: (a) delineating the evaluation context, (b) formulating evaluation questions, (c) data collection, (d) assessment and judgment, and (e) utilization of evaluation findings. As such, we account for a complete assessment of the role of citizens. Fischer (2002), for instance, stated that citizen involvement is quintessential during the entire research process, to ensure the validity of the findings produced (see also Juntti et al., 2009). As mentioned before, we do not argue for any particular role of citizen involvement. Again, our ambition is simply to take stock of the (theoretical) role given to citizens.

In what follows, we review how citizen involvement is conceived in the different models, per stage of the evaluation process. We focus on the general trends observed for effectiveness models, economic models and actor-oriented models as the three main types. Within the scope of the article, it is not possible to discuss all fine-grained nuances at the level of each of the 22 specific models, but we will refer to some clear examples per type to substantiate our point where relevant. In Table A1 in Appendix, the detailed findings of the document analysis can also be consulted.

Next to the review of the original texts of the models, we conducted an expert survey. These international experts were either the pioneers in introducing one or several of the 22 evaluation models, or—in case the founder of a particular evaluation model is no longer alive or not available—strongly contributed to its development. Hooghe et al. (2010, p. 692) suggest that expert surveys are appropriate if reliable information can be found with experts rather than in reliable documentary sources. The survey, launched between September and December 2019, listed the same factors that were included in the document analysis (see Table A2 in Appendix). About 60% of the evaluators responded (see Table A3 in Appendix for the overview), resulting in 12 completed surveys. While not all models are covered, the expert input helped us nuancing the results of the documentary analysis and supported us to assess whether the criteria of the document analysis allow adequate conclusions.

Let the Models Speak: Findings of the Document Analysis and the Expert Survey

Delineating the Evaluation Context—Are Citizens’ Perspectives Considered?

A quintessential part of an evaluation process is to delineate the context factors that provide the evaluation framework. According to Alkin and Jacobson (1983, p. 18), context refers to the elements or components of the situational backdrop against which the evaluation process unfolds. Examples of context factors are fiscal or other constraints on the evaluation, the social and political climate surrounding the project being evaluated, and the degree of autonomy which the project enjoys. In other words, the context concerns the framework within which the evaluation is

conducted (Alkin & Jacobson, 1983, p. 62). Mark and Henry (2004, p. 37) distinguish between the “resource context” that involves the human and other resources allocated to the evaluation and the “decision/policy context” that consists of the cultural, political and informational aspects of the organization involved in the policy implementation. Hence, delineating the evaluation context involves defining the rules of the game that an evaluator has to know prior to formulating the evaluation questions. As can be derived from the overview table, this broader evaluation context, and the role of citizens in particular, is differently reflected in the three main evaluation models, and in varying degrees of explicitness.

Starting with the effectiveness models, the picture is mixed. Some so-called theoretical approaches explicitly address the evaluation context, as it is at the center of their *raison d'être*. Stuffelbeam's CIPP model (2004)—which acronym refers to “context”—distinguishes four steps in their evaluation model: context, input, process, and product evaluation. Context, in this model, refers to the needs and associated problems linked to the evaluand. An evaluation can help identifying the necessary goals for improving the situation. Pawson and Tilley (1997)'s realistic evaluation revolves around a related notion of context: in their model social programs are conceived as influenced by the surrounding context. Some interventions might only work under certain circumstances. The same applies to logic models where citizens' perspectives can be a deeply influencing causal factor to realize a certain outcome. In attribute based approaches instead, citizens' perspective is seldom considered, precisely due to their research focus on the generalizability of results. Hansen (2005, p. 450) argues that program theory models are used to analyze the causal relations between context, mechanism, and outcome. Randomized-controlled trials (Campbell & Stanley, 1966) and quasi-experimental evaluation (Campbell & Stanley, 1966) particularly have the goal to control the context, and therefore cut out the perspective of citizens, in order to increase internal and external validity. As one expert phrased it: “In the best case, I was part of a [research team] that randomly sampled schools and that over-sampled based upon theoretically driven questions regarding the context. (...) But this may be the exception that proves the rule.”

Citizen perspectives are also only peripheral in economic models such as productivity and efficiency models. They have a strong focus on the relationship between invested resources and the output, and as such also apply a rather narrow or no perspective to citizens' considerations. Whether economic models take citizens' interests into account, as experts mentioned to us, is entirely in the hands of evaluators. For instance, cost-effectiveness models can provide an opportunity for stakeholders to reach consensus on what the most important objectives are, and how to measure them. The latter can be influenced by citizen's opinions, but most often it is not, also since citizens' response has to be translated into monetary values. Yet, we were told that evaluators can soften these conditions in order to give citizens a say in these evaluation models: “It is my belief that those factors that cannot be easily measured in monetary terms should continue to have a place in the analysis and this is perhaps where citizens can play a significant role.”

Both the document analysis and the experts confirmed that citizens' perspectives can best be accounted for in actor-oriented models. Models of democratic evaluation (House & Howe, 2000; MacDonald, 1976) and empowerment evaluation (Fetterman, 2001; Fetterman & Wandersman, 2007) are by nature conceived for citizens and stakeholders more in general. Both models' aim is to have an informed citizenry and community, in which the evaluator acts as a broker between different groups in society. For such societal groups, the evaluation can provide an opportunity to evaluate their performance and to accomplish their goals. In some cases, such as empowerment evaluation, citizens can even carry out the evaluation, with the support and facilitation of specialists. Moreover, citizens' perspective can serve as an important data resource, also in the design, interpretation, and promulgation stages. This is not to say that actor-oriented models always consider citizens' input, however. As for responsive evaluation models, for instance, it was stated “that there

is no expectation of citizens having a decision-making or participative role in any stage.” Hence, while the label “actor-oriented” may suggest this, it would be wrong to conclude that citizens have a stake per se in defining the framework within which the evaluation is conducted.

Evaluation Questions—Are Citizens’ Expectations Considered?

At the very essence of evaluations is the ambition to address particular questions about programs, processes, or products. Ideally, an evaluation provides information to a wide range of audiences that can be used to make better informed decisions, develop more appreciation and understanding, and gain insights for action. Evaluation questions typically consider the causal relationship between a program and an effect, but can also focus on the description of an object or a normative comparison between the actual and the desired state.

According to Preskill and Jones (2009), one way to ensure the impact of an evaluation is to develop a set of evaluation questions that reflect the expectations, experiences and insights of as many stakeholders as possible. They argue that all stakeholders are potential users of evaluation findings, which is why their input is essential to establish the focus and direction of the evaluation. Naturally citizens are important stakeholders that evaluators have to account for. When citizens’ expectations are already included at an early state of the evaluation, citizens’ specific information needs will most likely be addressed.

Again, and as expected, evaluation models vary strongly in respect of their focus of the evaluation questions and how they include citizens’ perspectives. Based on the review of original sources, effectiveness and economic models do not strongly consider citizen’s expectations. However, one should be careful in concluding that citizens have no role at all in this regard. According to some experts, citizens are increasingly having a say in practice in effectiveness models, as evaluators, “are driven by concerns with the ‘usefulness’ or ‘relevance’ of findings from RCTs for decision-making.” The same applies to economic models, where there seems to be an increased attempt to add evaluation questions that put the citizen in the process of placing monetary values on various policy goals. However, this is again not a principle but a choice by the evaluators: “Citizens’ perspectives can be considered (and often are) in articulating the theory of action and planned implementation and also in hypothesizing the resources required to operationalize that theory.” It should be pointed out that not all experts share this opinion though. For some, the top-down structure of economic evaluation models, being commonly decided by a funding body or other decision maker, makes citizen input far less important than other interests.

Notwithstanding these nuances, actor-oriented models are indeed much more oriented towards citizens’ needs. As explained by Hansen (2005, p. 449), stakeholder and client-oriented models particularly focus on the question whether stakeholders’ or clients’ needs are satisfied. Also in empowerment evaluations and participatory evaluations are citizens’ expectations usually the driving forces in developing evaluation questions. In participatory evaluation, the interests and concerns of citizens with least power are usually allocated some priority. Importantly, however, some experts argued that even these models cannot always include citizen’s interest: “Seldom in the past has 2% or more of the budget been warranted for inquiry into citizens’ expectations, experiences, and insights.” This observation itself is worth highlighting.

Evaluation Methods—Are Citizen’s Voices Considered?

In general, evaluators use a broad scope of methods to address the evaluation questions formulated. According to Widmer and De Rocchi (2012, pp. 98–100), evaluation methods tend to vary in four different aspects to regular research activities. To start with, evaluations frequently use a combination of methods and procedures. In doing so, one can distinguish between a method mix and triangulation.

Whereas the former refers to the use of different methods for answering different questions, the latter concerns the use of several different methods for answering the same question. Second, evaluations often deal with questions that focus on change, that is, whether a program has achieved a specific societal effect. However, these effects can often only be observed after a few months or even years. For this reason, comparisons in time axes, so-called before-and-after comparisons, are very common in evaluation. Third, the generalization of evaluation results is often only of secondary importance, mainly since the commissioner of an evaluation aims to understand the consequences of an intervention in a given context. Fourth and last, target/actual comparisons are frequently used in evaluations, which focus on the extent to which a program has achieved the declared objectives.

Theoretically speaking, citizens' voice can be integrated in the implementation of evaluation methods in several ways. Citizens' input can be directly used as a source for information, in particular in the context of triangulation. And where expert panels are used to assess the effects of an intervention, such results can in principle also be validated by ordinary citizens, with social interventions often explicitly aiming to change citizens' behavior. In sum, the perception of citizens can be a useful tool to observe how and to what degree an intervention has contributed to changes. As it was highlighted by one of the experts: "high quality evaluation design (see US program standards, American Evaluation Association) considers citizen stakeholder perspective in all aspects of the evaluation architecture, including evaluation methods." When reviewing the different models, however, the strong variability in considering citizen's voice is again apparent.

For economic models, such as cost-effectiveness analysis, it is not common to see citizen voice directly in data collection and data analysis. For cost-benefit analysis in particular it was mentioned that community activity does not always translate in monetary values. The same largely applies to effectiveness models, where the inclusion of citizens is rather exception than rule.

On the other side of the continuum, one can again identify the more actor-centered models, but the latter also have different traditions as to the involvement of citizens in data collection. For participatory evaluation, for instance, it will depend on the type of evaluation whether stakeholders will be actively involved, beyond providing input for the design and implementation of the evaluation. As it was mentioned to us: "the more technical these activities, the less likely it is for stakeholders to be active participants in data collection and analysis." In empowerment evaluation instead, it is part and parcel that citizens conduct the evaluation and determine what is a safe and appropriate method, in conjunction with a trained evaluator. The empowerment evaluator will not only make sure that citizens get the information they need through data collection, but will also help "with rigor and precision by helping them learn about the mechanics, ranging from data cleaning to eliminating compound questions in a survey." Also client-oriented models lend themselves generally well to the consideration of citizens' input. Utilization-oriented evaluation, for instance, is by design targeted at "identifying specific intended users for specific intended uses and engaging them interactively, responsively, and dynamically in all decisions, including methods decision." Whether these intended users are citizens, will then again depend on the particular evaluation or intervention at stake though. Altogether, actor-oriented models offer most potential to account for citizen input, but neither provide the guarantee that citizens' voices will be actively heard.

Assessment—Are Citizens' Values Considered?

To many, the very rationale of evaluations is to come to an assessment of public policies on the basis of criteria on which this judgment relies. According to Scriven (1991b: 91), "the term value judgment erroneously came to be thought of as the paradigm of evaluation claims." This theory of values can already be found in the early work by Rescher (1969: 72):

Evaluation consists in bringing together two things: (1) an object to be evaluated and (2) a valuation, providing the framework in terms of which evaluation can be made. The bringing together of these two is mediated by (3) a criterion of evaluation that embodies the standards in terms of which the standing of the object within the valuation framework is determined.

However, Shadish et al. (1991, p. 95) argue that both authors use different terms (Scriven standards; Rescher criterion) to explain the basis of judgements. Stockmann (2004, p. 2) illustrates that the assessment of the evaluated results is not anchored in given standards or parameters, but on evaluation criteria that can be very different. From his stance, evaluations are often meant to serve the benefit of an object, an action or a development process for certain persons in the groups. The evaluation criteria can be defined by the commissioner of an evaluation, by the target group, the stakeholders involved, by the evaluator him/herself or by all these actors together. It is obvious that the evaluation of the benefits by individual persons or groups can be very different, depending on the selection of criteria.

As Shadish et al. (1991) argue, evaluation criteria can be modeled along the values of stakeholders, which can potentially also include citizens. It is up to the evaluator to identify these values, to use them in constructing evaluation criteria and to conduct the evaluation in terms of those. Scriven (1986), however, rejects this model. He particularly considers it the role of evaluators to make appropriate normative, political, and philosophical judgments, as public interest is often too ambiguous to rely on.

As it turns out from the review, economic and efficiency models not explicitly consider citizens' values in the development of evaluation criteria, or only in an indirect way. As for effectiveness models, of which logic models are a case in point, it was mentioned that citizens values can be determinant in assessing the practical utility and related return of any policy, but whether this is indeed taken into account will depend on individual evaluators.

Also economic models have not explicitly built in citizen values, but only indirectly address these. For instance, in cost-utility analysis, "citizens' values are captured in the utility measure, but this is only in service to the overarching efficiency criterion"; or in cost-effectiveness analysis, citizen values are generally considered in analyzing distributional consequences or "who pays" and "who benefits."

Conversely, actor-oriented models hold more promises in considering citizens' input in establishing judgement criteria as the document analysis and the expert survey revealed. Empowerment evaluation is most outspoken in this respect, where community-based initiatives are evaluated against criteria that have been bottom-up derived, hence accounting for community values. As one of the experts explained: citizens "literally rate how well they think they are doing concerning the key activities that the group thinks they need to assess together (...) they also engage in a dialogue about their rating to ensure the assessment is meaningful to them." Participatory evaluation is another example, where citizen's input is a key concern in setting the criteria for what constitutes a "good" and "successful" program. As it was pointed out by one of the experts: "the point of view of program participants can be quite different from the 'expert' view." This exactly resonates with the legitimacy and epistemological motives that we highlighted in the introduction.

Utilization—Are Citizens' Interpretations of Findings Considered?

As pointed out by scholars as Eberli (2019, p. 25), evaluation utilization can be approached from multiple conceptual angles. While the terms "use" or "utilization" are applied synonymously for the use or application of evaluations (Alkin & Taut, 2002; Henry & Mark, 2003; Johnson et al., 2009), "utility" refers to the extent to which an evaluation is relevant to a corresponding question

(Leviton & Hughes, 1981). “Usefulness” on its part reflects the subjective evaluation of the quality and utility of an evaluation (Stockbauer, 2000, pp. 16–17), and refers to the use of evaluation results by the evaluated institution with concrete consequences following the evaluation findings. In evaluation literature, one can identify a large number of attempts to classify the use of systematically generated knowledge. One of the most common distinctions conceives three different types of use (Alkin & King, 2016): Instrumental use refers to the direct use of systematically generated knowledge to take action or make decisions. Conceptual use points at indirect use of systematically generated knowledge that opens up new ways of thinking and understanding, or that generates new attitudes or changes existing ones. In addition, one can distinguish symbolic use which refers to the use of systematically generated knowledge to support an already preconceived position. This in order to legitimize, justify, or convince others of this position. Noteworthy is also the concept of process use, which authors as Patton (2008, p. 90) added to the typology. Process use implies use occurring due to the evaluation process and not due to its results.

In recent times, scholars have discussed the role of evidence in public discourse and how citizens can be included in this context (Boswell, 2014; Pearce et al., 2014; Wesselink et al., 2014). They argue that the use of citizen information in policy discourse, and an interpretive view to policy making can lead to more reasoned debates. Schlaufer et al. (2018), for instance, empirically showed that the use of evaluations in direct-democratic debate provides policy-relevant information and increases interactivity of discourse. They also revealed that evaluations are particularly used by elite actors in their argumentation, de facto leading to a separation of the arena of deliberation between experts and general public. El-Wakil (2017), however, showed that a facultative referendum, one of the instruments that can be used in a direct democracy, gives political actors strong incentives to think in terms of acceptable justifications. Considering these contributions, evaluators might not only include citizens when assessing evaluation results, but also encourage evaluation clients to deliberate the use of evaluation with interested citizens.

Based on our analysis, it is clear that many models do not structurally involve citizens, generally speaking at least. We largely attribute this to the fact that debates about the use of evaluations only emerged after the development of many of the evaluation models. Take effectiveness models, such as RCTs for instance, where citizens are “rarely” involved in the interpretation, “except to include quotes or feedback directly on their experience using an intervention.” Also in efficiency models, the consideration of citizens is rather uncommon. The following expert quote for cost-benefit analysis (CBA) is illustrative: “this [the involvement of citizens in the interpretation of findings] is a critical question and sometimes ignored by CBA evaluators” (...) and “CBA often is less transparent on the assumptions that drove the analysis.” For cost-utility analysis, the expert indicated “I have never seen this done, though this is a principle I would support.”

On the other side of the spectrum, we can again situate actor-oriented models. Stakeholder-oriented models, such as empowerment evaluation, can be said to have most built in possibilities for citizen involvement. Empowerment evaluations serve in the first place the communities and stakeholders that are affected by a program. As they are the primary audience in evaluations, they can also be easily considered in the utilization stage. As one of the experts mentioned: citizens “self-assess by selecting the key activities they want to evaluate and then rate how well they are doing. Once they discuss what the ratings mean, they are prepared to move to the 3rd step and plan for the future (...). The plans for the future represent their strategic plan.” Also MacDonald’s democratic evaluation model which has particularly been designed for the sake of informing the community. The evaluator “recognizes the value pluralism and seeks to represent a range of interests in its issue formulation. His main activity is the collection of definitions and reaction to the program” (MacDonald, 1976, p. 224). This is not to say that all actor oriented models have guarantees for citizen involvement, and that all stakeholder oriented models incorporate citizen input by definition. Especially for responsive evaluation, we were given a less positive image by experts than what we derived from the reading of the

original models: “seldom have responsive evaluators or their clients been involved with formal procedures for citizen interpretation of the findings. The most common ending has been that the evaluand administrators or their superiors have moved beyond their concerns when the evaluation was commissioned, and have confidence that they can do what needs to be done.” Also client-oriented models not necessarily pay much attention to citizens, or apply a relative elitist perspective to the interpretation of the evaluation findings. Utilization focused evaluation, for instance, is solely oriented towards the intended users of an evaluation, most often program managers (Patton, 2008). And when citizens are involved in utilization-focused evaluation, this is mainly restricted to the “primary intended users.” In sum, actor oriented does not necessarily mean that there is also room for citizen input, also in the utilization stage of an evaluation cycle.

Conclusion

Do the different evaluations models consider citizens during the evaluation process? Whereas citizen participation in evaluation holds the potential to realize important democratic values, there is not much evidence as to whether and how existing evaluation models account for citizens’ considerations. To our knowledge, this article is the first that relies on a unique combination of document analysis and expert input in order to provide a review of the role of citizens in evaluation. The distinction between different dimensions of the evaluation process furthermore allows for a comprehensive and fine-grained perspective. Not surprisingly, actor-oriented models have most built-in guarantees for a relatively extensive involvement of the public and communities throughout the entire evaluation process, as these models have also been particularly conceived for these purposes. Yet, as our analysis revealed, substantial variation also exists across these different models, and across the different stages of the evaluation process. At the risk of crude generalization, it can be said that citizens’ input is usually relatively built in actor oriented models when considering “context,” “evaluation questions,” and “assessment,” but this is much less evident for the stages of “data collection and data analysis” and especially for the “utilization of findings.” Experts also put some nuance to what we derived from the mere reading of the original models in this respect.

By the same token, it would be wrong to assume that citizens’ considerations are completely ruled out in the other models, but less so explicitly, or dependent on the actual focus on the evaluation. Many models make explicit reference to “stakeholders” as broader denominator, but individual citizens are usually not targeted. True, many models were already developed before the heydays of collaborative governance or citizen science, which can explain why citizen involvement is not at the center of attention in many models.

Admittedly, many models have been revised over time. Since we have purposely focused on the original models instead of the newer developments, some models may have been assessed more negative or positive compared to later developments. For instance, Stufflebeam and Zhang (2017) discuss prominently the involvement of citizens in the updated CIPP evaluation model. In contrast, newer interpretations of democratic evaluation rather suggest to focus on macro-structures such as democracy institutions, democratic transitions, democratic renewal instead of citizens (Hanberger, 2018). Our deliberate choice to focus on the original models, however, is based on the firm belief that basic principles such as the inclusion of stakeholders can be deduced from the original models. Besides, given the richness of the evaluation field, it would not have been possible to address all latest developments of original models within the scope of this manuscript. We acknowledge though that this restriction is a limitation of this article.

This being said, existing literature on evaluation models only seldom distinguishes between stakeholders and citizens. By making this distinction explicit, we have attempted to make an important contribution to evaluation practice. More in particular, we would like to raise three issues that are important when planning to include citizens in an evaluation. At the beginning of the evaluation

project, evaluators should clearly define the purpose and the reason of the evaluation. Even though formative evaluations are more likely to be targeting citizens' input, this does not mean that the population cannot provide valuable feedback for summative evaluations. Our review provides a first orientation on available models that evaluators can choose from if they want to integrate citizens' interest.

Second, evaluators need to team up with commissioners of evaluations, and help them deciding on how to consider citizens' interest. The state transformation to modern technologies already provides public administrations various modalities for the involvement of citizens. However, public servants often are unsure how and to what degree they should engage citizens in the assessment of public policies. Therefore, evaluators need to support clients with the decision on how to consider citizens' interest. Typical citizens involvement includes community based practices where citizens not only accompany evaluation processes, but also deliberate about the content and findings of the evaluations.

Third, evaluators should reserve time for calibration and testing the evaluation models when planning the timeline of the evaluation. No matter what type of model is adopted, evaluators should always—in consultation with their clients—pretest the actual experience of real participants. This allows them to improve the involvement of citizens and find the right balance between valuable inclusion and infinite discussion. While this article assumes that citizen involvement in evaluation could be important for democratic values, whether and how the link between the two actually materializes in practice is an empirical issue to study. To be clear, it is not our ambition to uncritically call for increased citizen participation in evaluation at all costs. Evaluation is an inherently contingent undertaking, in which a plethora of contextual factors will determine whether it is wise to involve citizens. The nature of the policy problem may be one element to consider in this regard. One may state, for instance, that more technical problems likely lend itself less for citizen engagement. In the same vein, different civic epistemologies in various countries or policy sectoral settings (Jasanoff, 2011) can also make that citizen input will be differently valued. In fact, to take it further, the decision to involve citizens in evaluation is not merely a matter of model or design choice, or having the appropriate resources for it (Fung, 2015). The major decision is of political nature in the first place, with evaluation commissioners that should ultimately be willing to actively engage citizens in one or several stages of the evaluation process. With many evaluations being intrinsically political (Weiss, 1993), the decision whether and how to involve citizens is *de facto* also a decision about the allocation of power (Juntti et al., 2009).

Acknowledgments

Previous version of this paper was presented at the International Conference on Public Policy 2019 in Montreal, Canada, at the European Group for Public Administration Annual Conference 2019 in Belfast, Northern Ireland. The authors thank all the participants for their feedback. We are grateful to Sebastian Lemire as well as the three anonymous reviewers for their valuable remarks.


Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Pirmin Bundi  <https://orcid.org/0000-0001-5047-072X>

References

- Alkin, M. C. (2017). When is a theory a theory? A case example. *Evaluation and Program Planning, 63*, 141–142. doi:10.1016/j.evalprogplan.2016.10.001
- Alkin, M. C., & Jacobson, P. (1983). *Organizing for evaluation use. A workbook for administrators. evaluation productivity project*. Los Angeles: California University. Centre for the Study of Evaluation.
- Alkin, M. C., & King, J. A. (2016). The historical development of evaluation use. *American Journal of Evaluation, 37*(4), 568–579. doi:10.1177/10982140166665164
- Alkin, M. C., & Taut, S. M. (2002). Unbundling evaluation use. *Studies in Educational Evaluation, 29*, 1–12. doi:10.1016/S0191-491X(03)90001-0
- Alkin, M. C., & Vo, A. T. (2017). *Evaluation essentials: From A to Z*. New York and London: Guilford Publications.
- Bäckstrand, K. (2006). Multi-stakeholder partnerships for sustainable development: Rethinking legitimacy, accountability and effectiveness. *European Environment, 16*(5), 290–306. doi:10.1002/eet.425
- Boswell, J. (2014). ‘Hoisted with our own petard’: Evidence and democratic deliberation on obesity. *Policy Sciences, 47*(4), 345–365. doi:10.1007/s11077-014-9195-4
- Brandon, P. R., & Fukunaga, L. L. (2014). The state of the empirical research literature on stakeholder involvement in program evaluation. *American Journal of Evaluation, 35*(1), 26–44. doi:10.1177/1098214013503699
- Brunner, I., & Guzman, A. (1989). Participatory evaluation: A tool to assess projects and empower people. *New Directions for Evaluation, 1989*(42), 9–18. doi:10.1002/ev.1509
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin.
- Chen, H. T. (1994). *Theory-driven evaluations*. Newbury Park, CA: Sage.
- Chen, H. T., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review, 7*(3), 283–302. doi:10.1177/0193841X8300700301
- Coffman, J. (2007). What’s different about evaluating advocacy and policy change. *The Evaluation Exchange, 13*(1), 2–4.
- Contandriopoulos, D., & Brousselle, A. (2012). Evaluation models and evaluation use. *Evaluation, 18*(1), 61–77. doi:10.1177/1356389011430371
- Cousins, J. B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation. Kluwer international handbooks of education* (pp. 245–265). Dordrecht: Springer.
- Cousins, J. B., & Chouinard, J. A. (2012). *Participatory evaluation up close: An integration of research based knowledge*. Charlotte, North Carolina: IAP.
- Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14*(4), 397–418. doi:10.3102/01623737014004397
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation, 1998*(80), 5–23. doi:10.1002/ev.1114
- Eberli, D. (2019). *Die Nutzung von Evaluationen in den Schweizer Parlamenten*. Zürich: Seismo.
- El-Wakil, A. (2017). The deliberative potential of facultative referendums: Procedure and substance in direct democracy. *Democratic Theory, 4*(1), 59–78. doi:10.3167/dt.2017.040104
- Falanga, R., & Ferrão, J. (2021). The evaluation of citizen participation in policymaking: Insights from Portugal. *Evaluation and Program Planning, 84*, 101895. doi:10.1016/j.evalprogplan.2020.101895
- Fetterman, D. M. (2001). *Foundations of empowerment evaluation*. Thousand Oaks, CA: Sage.
- Fetterman, D. M., & Wandersman, A. (2007). Empowerment evaluation: Yesterday, today, and tomorrow. *American Journal of Evaluation, 28*(2), 179–198. doi:10.1177/1098214007301350
- Fischer, F. (2002). *Citizens, experts and the environment*. Durham: Duke University Press.
- French, M. T. (2000). Economic evaluation of alcohol treatment services. *Evaluation and Program Planning, 23*(1), 27–39. doi:10.1016/S0149-7189(99)00035-X

- Fung, A. (2015). Putting the public back into governance: The challenges of citizen participation and its future. *Public Administration Review*, 75(4), 513–522. doi:10.1111/puar.12361
- Greene, J. G. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, 12(2), 91–116. doi:10.1177/0193841X8801200201
- Guijt, I. (2014). *Participatory approaches, methodological briefs: Impact evaluation 5*. Florence: UNICEF Office of Research.
- Hanberger, A. (2001). What is the policy problem? Methodological challenges in policy evaluation. *Evaluation*, 7(1), 45–62. doi:10.1177/13563890122209513
- Hanberger, A. (2018). Rethinking democratic evaluation for a polarised and mediated society. *Evaluation*, 24(4), 382–399. doi:10.1177/1356389018802133
- Hansen, H. F. (2005). Choosing evaluation models: A discussion on evaluation design. *Evaluation*, 11(4), 447–462. doi:10.1177/1356389005060265
- Henry, G. T., & Mark, M. M. (2003). Toward an agenda for research on evaluation. *New Directions for Evaluation*, 2003(97), 69–80. doi:10.1002/ev.77
- Hilhorst, T., & Guijt, I. (2006). *Participatory monitoring and evaluation a process to support governance and empowerment at the local level. Guidance paper*. KIT: Amsterdam.
- Hooghe, L., Bakker, R., Brigevid, A., De Vries, C., Edwards, E., Marks, G., & Vachudova, M. (2010). Reliability and validity of measuring party positions: The chapel hill expert surveys of 2002 and 2006. *European Journal of Political Research*, 49(5), 687–703. doi:10.1111/j.1475-6765.2009.01912.x
- House, E. R., & Howe, K. R. (2000). Deliberative democratic evaluation in practice. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models. Evaluation in education and human services* (pp. 409–421). Dordrecht: Springer.
- Irwin, A. (1995). *Citizen science: A study of people, expertise, and sustainable development*. New York: Routledge.
- Jasanoff, S. (2011). *Designs on nature: Science and democracy in Europe and the United States*. Princeton: Princeton University Press.
- Johnson, K., Greenesid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, 30(3), 377–410. doi:10.1177/1098214009341660
- Juntti, M., Russel, D., & Turmpenny, J. (2009). Evidence, politics and power in public policy for the environment. *Environmental Science & Policy*, 12, 207–215. doi:10.1016/j.envsci.2008.12.007
- Kahane, D., Loptson, K., Herriman, J., & Hardy, M. (2013). Stakeholder and citizen roles in public deliberation. *Journal of Public Deliberation*, 9(2), 1–37.
- Kee, J. E. (1994). Benefit-cost analysis in program evaluation. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (pp. 456–488). San Francisco: Jossey-Bass.
- Kibukho, K. (2021). Mediating role of citizen empowerment in the relationship between participatory monitoring and evaluation and social sustainability. *Evaluation and Program Planning*, 85, 101911. doi:10.1016/j.evalprogplan.2021.101911
- Knowlton, L. W., & Phillips, C. C. (2012). *The logic model guidebook: Better strategies for great results*. Thousand Oaks, California: Sage.
- Levin, H. M., & McEwan, P. J. (2001). *Cost-effectiveness analysis: Methods and applications* (2nd edition). Thousand Oaks, CA: Sage.
- Leviton, L. C., & Hughes, E. F. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review*, 5(4), 525–548. doi:10.1177/0193841X8100500405
- Macdonald, B. (1976). Evaluation and the control of education. In D. Tawney (Ed.), *Curriculum evaluation today: Trends and implications* (pp. 125–136). London: MacMilland Education.
- Madaus, D. L., Stufflebeam, G. F., & Kellaghan, T. (2000). *Evaluation models: Viewpoints on educational and human services evaluation*. New York: Kluwer Academic Publishers.
- Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), 35–57. doi:10.1177/1356389004042326

- Mohr, L. B. (1995). *Impact analysis for program evaluation*. Thousand Oaks, California: Sage.
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th edition). Thousand Oaks, CA: Sage. Wholey.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Thousand Oaks, California: Sage.
- Pearce, W., Wesseling, A., & Colebatch, H. (2014). Evidence and meaning in policy making. *Evidence & Policy: A Journal of Research, Debate and Practice*, 10(2), 161–165. doi:10.1332/174426514X13990278142965
- Plottu, B., & Plottu, E. (2009). Approaches to participation in evaluation: Some conditions for implementation. *Evaluation*, 15(3), 343–359. doi:10.1177/1356389009106357
- Preskill, H., & Jones, N. (2009). *A practical guide for engaging stakeholders in developing evaluation questions*. Princeton: Robert Wood Johnson Foundation.
- Robinson, R. (1993). Cost-utility analysis. *British Medical Journal*, 307(6908), 859–862. doi:10.1136/bmj.307.6908.859
- Schlauffer, C., Stucki, I., & Sager, F. (2018). The political use of evidence and its contribution to democratic discourse. *Public Administration Review*, 78(4), 645–649.
- Schmidt, V. A. (2013). Democracy and legitimacy in the European Union revisited: Input, output and ‘throughput’. *Political Studies*, 61(1), 2–22. doi:10.1111/j.1467-9248.2012.00962.x
- Scriven, M. (1973). Goal-free evaluation. In E. R. House (Ed.), *School evaluation: The politics and process* (pp. 319–328). Berkeley: McCutchan.
- Scriven, M. (1986). New frontiers of evaluation. *Evaluation Practice*, 7(1), 7–44. doi:10.1177/109821408600700102
- Scriven, M. (1991a). Pros and cons about goal-free evaluation. *Evaluation Practice*, 12(1), 55–62. doi:10.1177/109821409101200108
- Scriven, M. (1991b). *Evaluation thesaurus*. Sage.
- Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Thousand Oaks, California: Sage.
- Sonnichsen, R. (1989). Advocacy evaluation: A strategy for organizational improvement. *Science Communication*, 10(4), 243–259.
- Stake, R. (2004). *Standards-based & responsive evaluation*. Thousand Oaks, California: University of Illinois at Urbana-Champaign, USA: California: Sage Publications.
- Stockbauer, U. (2000). *Was macht Evaluationen nützlich? Grundlagen und empirische Untersuchungen zum Thema Verwertung und Verwertbarkeit von Evaluationen* (Dissertation at the University of Salzburg, Salzburg).
- Stockmann, R. (2004). *Was ist eine gute evaluation? Einführung zu Funktionen und Methoden von evaluationsverfahren*. CEval-Arbeitspapier, 9. Saarbrücken: Universität des Saarlandes.
- Stufflebeam, D. (1983). The CIPP model for program evaluation. In G. F. Madaus, M. S. Scriven, & D. L. Stufflebeam (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 117–141). Boston: Kluwer- Nijhoff.
- Stufflebeam, D. (2001). Interdisciplinary Ph.D. Programming in evaluation. *American Journal of Evaluation*, 22(4), 445–455. doi:10.1177/109821400102200323
- Stufflebeam, D. L. (2004). The 21st century CIPP model. In M. C. Alkin (Ed.), *Evaluation roots. Tracing Theorists’ views and influences* (pp. 245–266). Thousand Oaks, California: Sage.
- Stufflebeam, D. L., & Zhang, G. (2017). *The CIPP evaluation model: How to evaluate for improvement and accountability*. Guilford Publications.
- Sturges, K. M., & Howley, C. (2017). Responsive meta-evaluation: A participatory approach to enhancing evaluation quality. *American Journal of Evaluation*, 38(1), 126–137. doi:10.1177/1098214016630405
- Taut, S. (2008). What have we learned about stakeholder involvement in program evaluation? *Studies in Educational Evaluation*, 34(4), 224–230. doi:10.1016/j.stueduc.2008.10.007
- Tyler, R. W. (1942). General statement on evaluation. *Journal of Educational Research*, 35, 492–501. doi:10.1080/00220671.1942.10881106

- Vedung, E. (1997). *Public policy and program evaluation*. New York: Routledge.
- Vedung, E. (2004). Evaluation models and the welfare sector. In I. Julkunen (Ed.), *Perspectives, model and methods in evaluating the welfare sector – A nordic approach* (pp. 41–52). Helsinki: National Research and Development Centre for Welfare and Health.
- Weiss, C. H. (1993). Where politics and evaluation research meet. *Evaluation Practice*, 14(1), 93–106. doi:10.1177/109821409301400119
- Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation*, 1997(76), 41–55. doi:10.1002/ev.1086
- Wesselink, A., Colebatch, H., & Pearce, W. (2014). Evidence and policy: Discourses, meanings and practices. *Policy Sciences*, 47(4), 339–344. doi:10.1007/s11077-014-9209-2
- Wholey, J. S. (1989). Introduction: How evaluation can improve agency and program performance. In: J. S. Wholey & K. E. Newcomer (eds). *Improving government performance: Evaluation strategies for strengthening public agencies and programs* (pp. 1–12). San Francisco: Jossey-Bass.
- Whitmore, E. (1998). Understanding and practicing participatory evaluation. *New Directions for Evaluation*, 80, 1–104. doi:10.1002/ev.1113
- Widmer, T., & De Rocchi, T. (2012). *Evaluation: Grundlagen, Ansätze und Anwendungen*. Zürich/Chur: Rüegger.

Appendix A

See Tables A1–A3.

Table A1. The Involvement of Citizens in Evaluation Models. Review of Original Sources.

	Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization	
Effectiveness models	Objective-based evaluation (Tyler, 1942)	Evaluator identifies situations in which citizens can be expected to display certain types of behavior in order to achieve objectives.	Questions related to attending objectives.	Depending on the evaluation object, citizens may be the sample for trials that provide evidence regarding each objective	The purpose of evaluation is to validate the program's hypotheses. Distinction between real observed objectives and perceived objectives. The perception of citizens can give indications about the latter	Evaluator develops means in order to interpret and use the results
Goal-free evaluation (Scriven, 1973, 1991a)	Citizens' consideration: *** Since the evaluator does not know program's objectives, the context has to be regarded. Citizens can be part of the context.	Citizens' consideration: Citizens' considerations related to the effects of a program.	Citizens' consideration: ** To ensure accuracy and bias control, the evaluator looks at co-causation and over determination as well as calls in a social process expert consultant to seek undesirable effects	Citizens' consideration: ** Determination of a single value judgment of the program's worth. Evaluators should use the needs of affected citizens as evaluation criteria	Citizens' consideration: * Goal-free evaluation is seen as a threat by many producers by Scriven, primarily because is less under the control of the program manager: "perhaps enough to prevent its use.	
	Citizens' consideration: **	Citizens' consideration: * Citizens' consideration: *	Citizens' consideration: * Citizens' consideration: **	Citizens' consideration: *** Citizens' consideration: **	Citizens' consideration: * Citizens' consideration: *	

(continued)

Table A1. Continued.

	Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization
Randomized-controlled trials (Campbell & Stanley, 1966)	Context is controlled by a random assignment of the treatment. Citizens' individual needs are not further considered.	Questions related to the causal effects of program.	Evaluators establish a control group, which are randomly assigned. Citizens can be members of groups, but do not have a specific role	Judgment only based on the determination of causal effects.	No information on how the results should be used
Quasi-experimental evaluation (Campbell & Stanley, 1966)	Citizens' consideration: * Context is not controlled by a random assignment, but a through a justified selection or matching. Individual needs should not be further considered.	Citizens' consideration: — Questions related to the causal effects of program.	Citizens' consideration: * Evaluators establish a control group, which are selected on the basis of certain criteria. Citizens can be members of groups, but do not have a specific role	Citizens' consideration: — Judgment only based on the determination of causal effects.	Citizens' consideration: — No information on how the results should be used
CIPP model (Stufflebeam, 1983, 2001, 2004)	Citizens' consideration: * An evaluator should conduct a "context evaluation" that serves to identify the needs and associated problems in the concerned context. This helps to define the objectives that	Citizens' consideration: — Evaluator engages with a "representative stakeholder panel to help define the evaluation questions, shape evaluation plans, review draft	Citizens' consideration: * CIPP model does not provide information on the data collection, since it is a rather schematic order for an evaluation	Citizens' consideration: — The evaluation criteria focus on the improvement of a program. The inclusion of as many stakeholders as possible should enable a broad scheme of	Citizens' consideration: — Rather formative alignment: Evaluations should provide a continual information stream to decision makers to ensure that programs continually improve

(continued)

Table A1. Continued.

	Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization
	are necessary to improve the situation. Everything should be considered that influences the object of evaluation, including citizens.	reports and disseminate the findings" This stakeholder panel is the primary group with whom the evaluator regularly interacts.		perspectives. However, the involvement of citizens tends to be limited, as they are often poorly organized	their services. Moreover, evaluation should help aid decision makers in allocating resources to programs that best serve clients. However, citizens not part of this process, elitist approach of use Citizens' consideration: *
	Citizens' consideration: *** Evaluator is problem-oriented and starts from the counterfactual (state without intervention). Citizens can to some extent be considered by including the perspective of citizens towards a problem.	Citizens' consideration: * The central aspect of the model is the causal reconstruction of the mode of action of a program. Focus of questions not only towards objectives, but also activities and subobjectives. Citizen expectation on a policy can hardly influence the objectives or activities.	Citizens' consideration: — Models focusses on causal mechanism, but is open towards qualitative research methods. However, citizen's involvement restricted to their participation in the activities.	Citizens' consideration: —* The evaluation criteria are based on a causal chain that reflects the program theory. The causal sequence consists of various subobjectives, which lead to the outcome of interest resp. the ultimate outcome. Achieving the latter is the most important criterion. Citizen involvement	No information on how the results should be used.
Impact analysis (Mohr, 1995)					

(continued)

Table A1. Continued.

Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization
<p>Theory-driven evaluation (Chen & Rossi, 1983; Chen, 1994; Weiss, 1997)</p> <p>Citizens' consideration: ** The model is concerned with identifying secondary effects and unintended consequences. Change model explains how under which conditions certain effects appear, while the action model reflects about who I addressed and should conduct the activities. The citizens' perspective can be integrated, but it is not in the center of interest</p>	<p>Citizens' consideration: * Evaluator should rely on social science theory for the identification of potential areas for investigation in order to identify effects that go beyond those goals. The theories that the model seeks to construct are "plausible and defensible models of how programs can be expected to work." Citizens expectations not relevant, in contrast: Social theories important to illustrate that certain expectations are unrealistic.</p>	<p>Citizens' consideration: * Model acknowledges paradigm that accepts experiments and quasi experiments as dominant research designs, but argues that these devices should be used in conjunction with a priori knowledge and theory to build models of the treatment process and implementation system to produce better evaluations. Open towards citizens involvement as part of implementation phase.</p>	<p>restricted to their significance to the ultimate outcome. Citizens' consideration: * No information on what evaluation criteria should be used. Citizens' consideration: — No information on how the results should be used.</p>	<p>Citizens' consideration: — No information on how the results should be used.</p>

(continued)

Table A1. Continued.

	Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization
Logic models (Knowlton & Phillips, 2012)	<p>Citizens' consideration: ** The model should be socially constructed, which means that they will inevitably reflect assumptions, expectation, use and other context features. Citizens might be part of this context.</p>	<p>Citizens' consideration: * Questions should combine the short and long-term effects of a program with its activities and the underlying theoretical assumptions. Expectations from all actors are important, as well citizen.</p>	<p>Citizens' consideration: ** Method to visualize the planned activities and the expected results of a program without making a statement about causal relationships. Citizen can in general participate in those activities.</p>	<p>Citizens' consideration: — No information on what evaluation criteria should be used.</p>	<p>Citizens' consideration: — No information on how the results should be used.</p>
	<p>Citizens' consideration: ** The model does not only want to provide a causal effect, but also wants to find out under which circumstances a program works. Hence, a program depends highly on the context. The perspective of citizen is an important aspect of this context.</p>	<p>Citizens' consideration: ** The evaluator has to answer which mechanisms are activated by a program and which circumstances are necessary in order to activate these mechanisms. The citizens' expectations can be to some extent</p>	<p>Citizens' consideration: ** Ideally, blind selection of intervention and control. If this is not possible, the selection criteria should be indicated. Citizens are not necessarily part of investigation.</p>	<p>Citizens' consideration: — Assessment criteria is based on the program theory and whether how it worked has been supported or refuted by analysis. Citizens' value is only relevant if it is an important aspect in the program theory.</p>	<p>Citizens' consideration: — Evaluation has the task of checking out rival explanations (i.e. adjudication), which then provides justification for taking one course of action rather than another (i.e. politics). Hence, evaluations might bring power to the decisions in decision-making.</p>

(continued)

Table A1. Continued.

	Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization	
Economic models	Cost-productivity model (Vedung, 1997)	<p>Citizens' consideration: *** The context of a program is not in the center of this model. Hence, the perspective of citizens can hardly be integrated.</p>	<p>Citizens' consideration: ** Evaluation questions are focused on the relationship between the invested cost and the outputs. The interpretation of this relationship can be based on the expectation of citizens.</p>	<p>Citizens' consideration: * Data collection relies strongly on the invested resources and the relationship to the output. Citizens' are less involved in the data collection.</p>	<p>Citizens' consideration: ** Most important criterion is the good economic performance, meaning the relationship between output of products and the resources invested. Citizens' values not considered.</p>	<p>Citizens can be included in this setting in order to decide about policy outcomes. Citizens' consideration: ** No information on how the results should be used.</p>
	Work-productivity model (Vedung, 1997)	<p>Citizens' consideration: * The context of a program is not in the center of this model. Hence, the perspective of citizens can hardly be integrated.</p>	<p>Citizens' consideration: ** Evaluation questions are focused on the relationship between the invested work and the outputs. The interpretation of this relationship can be based on the expectation of citizens.</p>	<p>Citizens' consideration: * Data collection relies strongly on the invested work and the relationship to the output. Citizens' are less involved in the data collection.</p>	<p>Citizens' consideration: * Most important criterion is the good economic performance, meaning the relationship between the invested work and the resources invested. Citizens' values not considered.</p>	<p>Citizens' consideration: — No information on how the results should be used.</p>

(continued)

Table A1. Continued.

	Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization
Cost-effectiveness analysis (CEA) (Levin & McEwan, 2001)	Citizens' consideration: *	Citizens' consideration: **	Citizens' consideration: *	Citizens' consideration: *	Citizens' consideration: —
	The aim of CEA is to compare the relative costs to the outcome of several courses of action. The analysis ideally takes a broad view of costs and benefits to reflect all stakeholders. Citizens can be part of them, but this is not a central characteristic of the method.	The model holds the potential to consider costs and benefits of all stakeholders, as long as a value can be put on these. Citizens can be one of the stakeholder groups considered, though this is not a central characteristic of the method.	The raw data for the CEA can come from a wide variety of sources, and different groups of stakeholders. Citizens can be one of these actors, but this is not a prerequisite.	Citizens' consideration: * CEA is meant to highlight preferences of different categories of stakeholders or actors involved in the intervention is planned. With citizens being often one of the affected parties, their values might often be considered. But this is not a prerequisite.	Citizens' consideration: — The model does not explicitly foresee a deliberation of evaluation results by citizens.
Cost-Benefit analysis (CBA) (Kee, 1994)	Citizens' consideration: *	Citizens' consideration: *	Citizens' consideration: *	Citizens' consideration: *	Citizens' consideration: —
	The aim of CBA is to compare the total costs of an intervention with its benefits, by using monetary units. The best cost-benefit analyses take a broad view of costs and benefits, reflecting the interests of all stakeholders who	The model holds the potential to consider costs and benefits of all stakeholders, as long as these can be captured in monetary values. Citizens can be one of the stakeholder groups	The raw data for the CBA can come from a wide a variety of sources, and different groups of stakeholders. Citizens can be one of these actors, but this is not prescribed by the model.	Citizens' consideration: * CBA is meant to highlight preferences of different categories of stakeholders or actors involved in the intervention is planned. With citizens being often one of the affected parties, their	Citizens' consideration: — The model does not explicitly foresee a deliberation of evaluation results by citizens.

(continued)

Table A1. Continued.

	Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization
	will be affected by the program. Citizens can be part of them, but this is not a central characteristic of the method.	considered, though this is not a central characteristic of the method.		values might often be considered. But this is not a prerequisite.	
	Citizens' consideration: *	Citizens' consideration: *	Citizens' considerations: *	Citizens' considerations: *	Citizens' consideration: —
Cost-minimization analysis (CMA) (French, 2000)	The aim of CMA is to compare the costs of alternative interventions, which are assumed to have an equivalent effect. The model relies on a preceding evaluation of equivalence, in which different perspectives can be given priority (including citizens' perspectives). This is not a core part of the model, though.	Evaluation questions can involve the comparison of interventions, with an equivalent effect for citizens. However, this is not a core part of the model.	Information about calculating the costs of certain interventions can be collected among citizens, but this is not prescribed by the model.	The model can consider citizens' perspectives in the assessment of the costs of the intervention, but this is not explicitly prescribed.	The model does not explicitly foresee a deliberation of evaluation results by citizens.
	Citizens' consideration: *	Citizens' consideration: *	Citizens' considerations: *	Citizens' considerations: *	Citizens' consideration: —
Cost-utility analysis (CUA) (Robinson, 1993)	CUA compares the costs of different interventions with their outcomes	Evaluation questions focus on a person's level of well-being, but	CUA does not prescribe the involvement of	Apart from an individual person's level of wellbeing, other citizens'	The model does not explicitly foresee a deliberation of

(continued)

Table A1. Continued.

Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization
<p>measured in “utility based” units—that is, units that relate to a person’s level of wellbeing (often measured in terms of quality adjusted life year). The model applies a rather narrow perspective to citizens’ considerations.</p>	<p>does not consider other citizens’ expectations.</p>	<p>citizens in data collection</p>	<p>values are not considered.</p>	<p>evaluation results by citizens.</p>
<p>Citizens’ considerations: * Utilization-focused evaluation is anchored in the idea that evaluations should be planned and conducted in ways that enhance the likely utilization of both the findings and of the process itself to inform decision. The model requires identifying primary intended users, which guide the evaluation process. Primary intended users can</p>	<p>Citizens’ considerations: * Evaluation questions reflect the interests of the primary intended users. These can be citizens, if they indeed have the capacity to use the findings.</p>	<p>Citizens’ considerations: — Utilization-focused evaluation is not restricted to a particular type of evaluation method or data collection. It does not actively consider the involvement of citizens in data collection.</p>	<p>Citizens’ considerations: * The model can consider citizens’ values in the evaluation criteria, if the primary intended stakeholders consider this of importance.</p>	<p>Citizens’ considerations: — The model usually involves a discussion of the evaluation results with primary intended users. These can be citizens, if they indeed have the capacity to use the findings. But this is not necessarily the case.</p>
<p>Actor-oriented models Utilization-focused evaluation (Patton, 2008)</p>				

(continued)

Table A1. Continued.

	Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization
Management-oriented evaluation (Wholey, 1989)	<p>be citizens, in principle, if they indeed have the capacity to use the findings.</p> <p>Citizens' considerations: * Evaluators will collect the data about different decision alternatives that are relevant to managerial decision makers. The citizens' perspective is in principle not considered.</p>	<p>Citizens' considerations: * The evaluation will be designed to serve the informational needs and evaluation questions of managerial decision makers. There is in principle no room for citizen input.</p>	<p>Citizens' considerations: - Evaluators and program managers will work closely together to identify the decisions that should be made and the information that is needed for this. Citizens are not involved in data collection.</p>	<p>Citizens' considerations: * The judgement will be based on organizational values identified in interaction with managerial decision makers. Citizens' values are in principle not taken into account.</p>	<p>Citizens' considerations: * The evaluation is targeted to decision makers, with whom the results will be discussed. Citizens are not part of this audience.</p>
	<p>Citizens' considerations: — In advocacy evaluation, information is produced for the purpose of improving decision making and assisting organizations in the managing process. An organization's managerial perspective will drive</p>	<p>Citizens' considerations: * The evaluation will be designed to serve the informational needs and evaluation questions of managerial decision makers. There is in principle no room for citizen input.</p>	<p>Citizens' considerations: * Evaluators and program managers will closely work together to identify the data sources needed. The evaluator should especially consider the political nature of the decision-making</p>	<p>Citizens' considerations: — The judgement will be based on organizational values identified in interaction with managerial decision makers. Citizens' values are in principle not taken into account.</p>	<p>Citizens' considerations: — Both the evaluator and the client/host organization take an activist change agent role in the discussion of the evaluation findings. Citizens do not form part of the audience to which</p>

(continued)

Table A1. Continued.

	Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization
	the evaluation. The citizens' perspective is not considered.	principle no room for citizen input.	process and the effects evaluation activities can produce. Citizens are not involved in data collection.		results are communicated.
	Citizens' considerations: —	Citizens' considerations: —	Citizens' considerations: *	Citizens' considerations: —	Citizens' considerations: —
Responsive evaluation (Stake, 2004)	calls for evaluators being responsive to the information needs of various stakeholders, and will try to unravel how the program looks like to different people. Whether individual citizens are also be considered, will depend on the client commissioning the evaluation, and the focus of the evaluation.	The model clearly entails the possibility to ask for the citizen's perspective. It is particularly the task of the evaluator to identify different value perspectives to a program.	The choice of data-gathering activities is made in interaction with various groups having an interest in the program. If relevant for the intervention, this will also imply the involvement of individual citizens.	In responsive evaluation, the different value perspectives are referred to in reporting the success of a program or policy. Expressions of worth are gathered from different points of view.	Audience members are actively asked to react to the relevance of the findings, and will use different media accessible to them. A final report can be written, but this is not always the case. To the extent that citizens are considered a relevant party, citizens will also be informed about the evaluation findings.
	Citizens' consideration: **	Citizens' considerations: ***	Citizens' considerations: **	Citizens' considerations: ***	Citizens' considerations: **
Democratic evaluation (MacDonald, 1976; House & Howe, 2000)	The aim of democratic evaluation is to have an informed citizenry and	In democratic evaluation, the interests and values of major	Democratic evaluation tends to rely on inclusive data collection	In democratic evaluation, conclusions will be the result of	Democratic evaluation provides information to the community about a

(continued)

Table A1. Continued.

	Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization
	community, in which the evaluator acts as broker between different groups in society.	stakeholders involved in the intervention should be considered in the design of the evaluation.	practices which foster participation and collaboration, and in which community members and citizens have a key role. Dialogue and deliberation among stakeholders are key to arrive at conclusions.	extensive deliberations among stakeholder groups. All relevant community values need to be considered.	particular policy, and aims to serve all stakeholders.
	Citizens' considerations: ***	Citizens' considerations: ***	Citizens' considerations: ***	Citizens' considerations: ***	Citizens' considerations: ***
Participatory evaluation (Cousins & Whitmore, 1998); Guijt, 2014; Brunner and Guzman, 1989; Cousins/Whitmore, 1998)	Participatory evaluation involves the stakeholders of a policy intervention in the evaluation process. This involvement can concern one or more stages of the evaluation process. Stakeholders in participatory evaluation are commonly the people affected by the intervention.	The model clearly entails the possibility to ask about citizens' expectations, to identify locally relevant evaluation questions. Participatory evaluation can also be restricted to other stages of the evaluation process	Participatory evaluation encompasses a wide variety of participatory data collection methods in which citizens can play an active role. Stakeholder involvement can also concern other stages of the evaluation process.	Participatory evaluation provides the potential to reflect citizens' values, but stakeholder involvement can also be restricted to other stages of the evaluation process.	Participatory evaluation is designed to increase the ownership of evaluation findings, and to foster usage (amongst other purposes). Evaluation findings can be reported to citizens, but stakeholder involvement can also concern other stages of the evaluation process.

(continued)

Table A1. Continued.

	Context	Evaluation questions	Data collection	Assessment/ Judgment	Utilization
Empowerment evaluation (Fetterman, 2001; Fetterman & Wandersman, 2007)	<p>Citizens' considerations: ** Empowerment evaluation is a stakeholder involvement approach, and relies on the idea to help groups and communities evaluating their own performance and accomplish their goals. The model is particularly designed to evaluate community-based initiatives, but is not restricted to such audience.</p> <p>Citizens' considerations: ***</p>	<p>Citizens' considerations: ** The model is by nature based on the idea to capture communities' and stakeholder expectations. Citizens can be part of such groups. These constitute the basis for the evaluation process. Empowerment evaluation holds the potential to involve</p> <p>Citizens' considerations: ***</p>	<p>Citizens' considerations: ** Empowerment evaluation is collaborative and participative by nature, and explicitly relies on the involvement of citizens and community groups in data collection</p> <p>Citizens' consideration: ***</p>	<p>Citizens' considerations: ** Community-based initiatives are evaluated against criteria that have been bottom-up derived. Community values are hence fully respected. Such community values can also encompass individual citizens' values.</p> <p>Citizens' considerations: ***</p>	<p>Citizens' considerations: ** Evaluation findings in empowerment first place the communities and stakeholders concerned. They are, by nature, the primary audience of the evaluation.</p> <p>Citizens' considerations: ***</p>

Note for citizens' consideration, —: no specification; *: no or rather no consideration; **: some considerations; ***: strong consideration.

Table A2. Expert Survey Questionnaire.**Context**

The evaluation context defines the rules of the game that an evaluator has to know in order to evaluate a policy. Mark and Henry (2004) distinguish between the “resource context” that involves the human and other resources allocated to the evaluation and the “decision/policy context” that consists of the cultural, political and informational aspects of the organization involved in the policy implementation.

Considering this approach to “context,” could you please indicate whether and how the citizens’ perspective is considered in model * If not, please explain why?**

Evaluation questions

At the very essence of evaluations is the ambition to address particular questions about programs, processes, products. Evaluation questions not only focus on the causal relationship between a program and an effect, but can also focus on the description of an object or a normative comparison between the actual and the desired state.

Can you please indicate whether and how citizens’ expectations, experiences and insights are usually considered in developing evaluation questions in evaluations applying the model * If not, please explain why?**

Evaluation methods

Evaluators can use a broad scope of methods to address the evaluation questions formulated. Such methods usually concern data collection methods and techniques for data analysis.

Can you please indicate whether and how citizens’ voice is usually considered in (a) the selection of evaluation methods; (b) data collection and data analysis in evaluations applying the * model? If not, please explain why?**

Assessment

To many, the very rationale of evaluations is to come to an assessment of public policies on the basis of criteria on which the judgement relies.

Can you please indicate whether and how citizens’ values are usually considered in the development of evaluation criteria as for evaluations applying the *model? If not, please explain why?**

Utilization of evaluation results

An evaluation can serve multiple purposes, and evaluation results can be discussed with multiple actors.

Can you please indicate whether and how citizens are usually involved in the interpretation of the evaluation findings, in evaluations applying the *model? If not, please explain why?**

Table A3. List of Experts (in Alphabetical Order).

Name	Affiliation
David Fetterman	Fetterman & Associates
Jennifer Greene	University of Illinois
Ernest R. House	University of Colorado
James Edwin Kee	George Washington University
Julian King	Julian King & Associates
Patrick McEwan	Wellesley College
Michael Q. Patton	University of Minnesota
Robert Shand	American University
Robert E. Stake	University of Illinois
Elizabeth Tipton	Northwestern University
Evert Vedung	Uppsala Universitet
Lisa Wyatt Knowlton	Wyatt Advisors