Integrative Evaluation
Osvaldo Feinstein¹

Abstract

“Integrative evaluation” is an approach with two main phases: identification of plausible rival hypotheses and integration of rival hypotheses. The first phase may correspond to traditional adversary evaluation, whereas the second phase, that is not included in adversary evaluation, requires integrative thinking which can be applied when they are compatible and complementary. Integrative evaluation may facilitate “evaluative thinking” and contribute to deliberative or integrative democracy.

Keywords
adversary evaluation, synthesis, bias, integrative thinking.

This article is inspired by Robert Picciotto’s “Is Adversary Evaluation Worth a Second Look?” (Picciotto, 2019) which raised the important issue of how to make use of contrasting views of the impacts and value of the programs and policies that are evaluated. Whereas adversarial evaluation focused on competing accounts that are considered prior to making decisions, another tradition in evaluation has been the deliberate use of different

¹ Magister en Evaluación de Programas y Políticas Públicas, Universidad Complutense de Madrid, Madrid, Spain

Corresponding Author:
Osvaldo Feinstein, Magister en Evaluación de Programas y Políticas Públicas, CSEG, Edificio B. Mas Ferré, Campus Somosaguas, Madrid 28223, Spain
Email: ofeinstein@yahoo.com
perspectives as a way of increasing the depth of our understanding, like participatory evaluation. Engaging stakeholders at different phases of the evaluation can add value by providing different perspectives on what will be considered a credible, high quality and useful evaluation contributing to the program logic and framing of key evaluation questions. This article shows that, under certain conditions, adversary evaluation can be reframed as “integrative evaluation”, an approach that may facilitate “evaluative thinking” and contribute to deliberative or integrative democracy, mitigating discourse polarization. Furthermore, this approach can also be applied to cases when the rival hypotheses, points of view, perspectives or approaches are not adversarial.

On Adversary Evaluation

A useful definition of “Adversary-oriented Evaluation” is provided by Worthen et.al. (1997, p. 139): “a planned effort to generate opposing points of view within the overall evaluation”. It is to be noted that in this approach an effort is made not to avoid bias but to encourage both a positive and a negative bias, to shed light both on strengths and weaknesses of the program or intervention evaluated.

Adversary evaluation, bias blind spots and devil’s advocate

In Picciotto’s article, which deviates from the usual path to embrace the latest fashion, it is argued that "controversial social programs tend to elicit sharply different evaluative perspectives". See (Ortega y Gasset, 1938) and (Sharkey, 1970)

2 The reference to “perspectives” in this approach evokes the Spanish philosopher José Ortega y Gasset’s “perspectivism”. See (Ortega y Gasset, 1938) and (Sharkey, 1970)

3 https://www.betterevaluation.org/en/rainbow_framework/manage/understand_engage_stakeholders

4 Prof. George Julnes (personal communication) directed my attention to “theory knitting”. There is indeed a similarity with integrative evaluation, but there is a difference in the “objects” of the knitting: in integrative evaluation what are “knitted” are not theories but hypotheses, perspectives or approaches. Furthermore, whereas “In theory knitting, one attempts to integrate previous theories into a single higher order theory” Leeuw and Donaldson (2015, p. 474), integrative evaluation tries to integrate adversarial or non-adversarial hypotheses, perspectives or approaches rather than aiming at the development of a higher order theory.
judgments shaped by diverse stakeholders’ interests and evaluator’s dispositions. In such cases, it is not uncommon for two or more evaluations to be commissioned in order to accommodate different value frameworks and evaluation methods, a triangulation process that illuminates diverse facets of the evaluand and facilitates mediation of competing interests” (Picciotto, 2019. p.95)

What is crucial is to consider different perspectives, and this could be done even within a single evaluation, focusing first on achievements (intended and unintended), then on shortcomings (or failures or limitations), and finally making a judgement taking into account the achievements and shortcomings. A practical way to proceed is ensuring that the scope of work (or terms of reference) directs the attention of the evaluator(s) to both achievements and shortcomings. This modest approach to adversary evaluation (AE) may be more feasible than a full-fledged AE.\textsuperscript{5} It should be noted that the thrust of R.Picciotto’s article is indeed that the judicial evaluation model is usually too demanding and that it is only one embodiment of AE.

An alternative view is that cognitive biases are so pernicious that it is difficult for any of us to fully “step outside” our own perspective; there is a sort of “bias blind spot”. Given that a single evaluation team is subject to similar biases (e.g. they all receive funding from the same source, or have the same institutional incentives) even if they assign some members to play devil’s advocate, the resulting perspectives will still not be sufficiently diverse to eliminate bias. So it might be necessary to bring in individuals or teams with fundamentally different perspectives to get the full epistemic benefit of error canceling/triangulation. In fact it has been pointed out that there are studies which “suggest that individuals see the existence and

\textsuperscript{5} (Posner,1999) shows the usefulness of the adversary approach as a means to gather evidence.
operation of cognitive and motivational biases much more in others than in themselves(...) an asymmetry in perceptions of bias (...) blind spots” Pronin et.al (2002,p. 369), concluding that “Adversaries would be well advised to engage the efforts of sophisticated third parties. They should do so (...)in the hope that the third parties will help them in the search for common ground and for a future that both sides would find better than the status quo” (Pronin et.al, 2002,p. 380). Nevertheless it could be argued that in the case of a single evaluation team one of the roles that its team leader should play is that of a sophisticated third party.

Finally, it is interesting to note that a kind of Adversary Evaluation approach has been used for centuries by the Roman Catholic Church when considering someone for sainthood: it has a “devil’s advocate” to state the other side of the candidate’s story (Ellerman, 2004, p.166). In fact, several organizations use some form of devil’s advocacy (Ivancevich, Konopaske, & Matteson, 2011); for example, before making a major decision, such as entering a market or building a plant, Anheuser-Busch assigns some group the role of critic with the purpose of uncovering all possible problems with a particular proposal. This approach could be used in the context of a formative evaluation to assess and enhance the quality of proposed improvements in programs (or other type of interventions).

**Adversary Evaluation and Integrative Thinking**

The kind of adversary evaluation approach presented in R.Picciotto’s article, combined with the points made in the preceding paragraphs, could be framed in terms of an approach with two phases: the first one identifies plausible rival hypothesis, whereas the second phase attempts to integrate these hypotheses. Note that when one of the hypotheses or perspectives focuses on positive results, and another perspective highlights negative results, the

---

6 This alternative view was proposed by Prof. G.Wright (who also provided the reference to the article by Pronin et.al) in a personal communication.

7 David Ellerman discusses the general case for devil’s advocacy, identifying several classic antecedents such as Milton, Jefferson and John Stuart Mill. He also discusses problems in implementing devil’s advocacy (Ellerman, 2005,pp.170-1)
combination of both avoids the "positive bias", which neglects negative results (or shortcomings) and the "negative bias", which neglects positive results (or achievements).

The positive bias focuses the evaluation’s explanation on some factor(s) leaving in the shadow some other factor(s), which become the focus of the evaluation’s explanation triggered by a negative bias. The integrative challenge is to combine these apparently opposite explanations or perspectives, and this could be done in cases in which those explanations are compatible and complementary. Thus “integrative evaluation” can be perceived as an approach that promotes impartiality through explicit and compensated partialities.

The first phase could be seen as “deconstructing” or decomposing the evaluated intervention in terms of positive and negative results, which correspond to benefits and costs. In this phase assessments should be made of the worth and merit of the intervention, taking into account its relevance and sustainability (two criteria associated with worth, “doing the right things”) as well as its effectiveness (evaluation criterion associated with merit, “doing things right).

The second phase, would be fully evaluative, judging the value of the intervention in light of the evidence and judgements provided by the first phase, focusing on the impact of the intervention. The integration is a phase of the evaluation in which creative thinking should play an important role, “putting the data together in new ways to see the interactions among separate findings more holistically; synthesizing diverse themes in a search for coherence and essence” (Patton, 2018,p.22). As shown by Richard Florida, diversity encourages creativity (Florida,2002), so diversity in the composition of the evaluation team may be a way to encourage creativity through the integration of different perspectives.
Table 1: Phases of an Integrative Evaluation

First Phase:

Identify rival plausible hypothesis or alternative perspectives, searching for

  i) evidence of positive results/achievements of the intervention
  ii) evidence of negative results/failures of the intervention

Second Phase:

Consider the possibility of integrating the alternative perspectives

assess the value of the intervention taking into account the evidence gathered in the previous phase

When the alternative perspectives correspond to opposite views, which may be likened to a “thesis” and an “antithesis”, and in this case the integrative evaluation would be using a sort of “dialectical reasoning”: a process of “thesis” followed by “antithesis” – a proposition followed by a potential contradiction of that proposition⁸, followed by a synthesis that resolves the contradiction (Nisbett, 2015,p.206⁹). This is a kind of “heuristic dialectic” rather than an ontological or logical dialectic; the last one corresponds to the Hegelian approach, which is NOT followed in this article, whereas dialectics as a heuristic device is a pragmatic approach (Bhaskar, 2008) and (Wan 2013). The synthesis ¹⁰ can be considered an application

---

⁸ Albert Hirschman’s “self-subversion” is a kind of antithesis; Hirschman (1995). A Hirschmanian approach to development evaluation is discussed in (Feinstein,2018).

⁹ Nisbett’s presentation of dialectical reasoning is framed in practical terms without philosophical jargon and fits well with the purpose of this article in showing the role of synthesis or integration in going beyond adversarial positions.

¹⁰ “Synthesis in social science is the process of collating a number of arguments by extracting their essential elements and recombining them as a 'higher level' analysis that incorporates and accounts for these elements” (Harvey, 2019).
of “integrative thinking”¹¹, as shown in the following example from actual evaluations of rural development projects:

Evaluation (A) argues that the increase in the income of small farmers in region X was due to the availability of credit for small farmers, whereas evaluation (B) identifies the availability of a suitable technical package as the cause of the farmers increased income (i.e., the effect or outcome). An integrative alternative (synthesis) would be to consider both the availability of credit for small farmers & the availability of a suitable technical package as necessary conditions for increase in the farmers income. Neither of the two would have been sufficient to explain the farmers increased income, whereas the combination of the two is necessary and sufficient. So A and B can be integrated, becoming components of a synthesis (C).

The three views can be presented in a way that could allow for an empirical test:
A. Increase in credit plays by far the largest causal role in effecting the increase in farmers incomes, and this is largely independent of the availability of technical packages.

B. Availability of technical packages plays by far the largest causal role in effecting the increase in farmers incomes, and this is largely independent of the increase in credit.

C. The interaction of credit increase and availability of technical packages plays a much larger causal role than either of those factors do in isolation, in effecting the increase in farmers income.

Studies should be able to indicate which of those three views is more probable, and then once that becomes clear, further theories about why this is so could be generated, and in turn those

---

¹¹ For a discussion of “Integrative Thinking” in the context of management and leadership see (Martin, 2007)
theories could be tested against each other. If (C) proves most probable, it would represent an integrative approach to causal explanation\textsuperscript{12}.

M.Barker recently showed in the context of a discussion about species cohesion that an integrative option is an alternative that is worthwhile to consider when two approaches have concepts or causal explanations that can be compatible and complementary (Barker, 2019, p.18), rather than competing and irreconcilable accounts of reality. Following this line of argument, if it can be shown that /when/ the adversary evaluations are compatible and complementary then a synthesis or an integration would be feasible. Otherwise, an integrative evaluation could not be performed (an evaluability assessment could include an exploration of the feasibility of conducting an integrative evaluation).

A different way of presenting the previous example is by pointing out that X1 (credit) and X2 (technical packages) might both impact Y (farmers incomes) so a regression could be run on Y controlling for X1 and X2, including an interaction between X1 and X2, to account for the fact that the impact of X1 on Y depends on X2 (and vice versa). But the claim “X1 impacts Y” and the claim “X2 impacts Y” may not be perceived as antitheses of one another. So a better way of framing the argument would be:

Perspective a (Thesis): X1 impacts Y
Perspective b (Antithesis): X1 does not impact Y
Integration (Synthesis): X1 impacts Y in some cases (e.g. when X2 is high) but does not impact Y in other cases (when X2 is low)\textsuperscript{13}.

Another example would be the evaluation of economic and social policies in an open economy in which exports play an important role. The facts may be pointing out to positive

\textsuperscript{12} I owe this formulation to prof. Matthew Barker.
\textsuperscript{13} Graham Wright suggested this way of framing the example.
results in terms of poverty alleviation and economic growth. One perspective could emphasize the incentives introduced by the policies to promote production and transferences to the poor. A different perspective may emphasize the increased demand for the country’s exports (a commodity boom). But it could be the case that both explanations can be integrated in a single one which acknowledges the role of endogenous and exogenous factors. Furthermore, there could be different perspectives with respect to the poverty alleviation effects of transferences to the poor: a short-term perspective will focus on the short-term effect, which may be positive but not sustainable, which would be highlighted by a longer-term perspective. Both can be integrated, considering short and long-term effects, which are different time-levels or time-horizons. It is worthwhile to note that integration could also be done among hypotheses, approaches and/or perspectives corresponding to other different types of levels, such as macro, meso and micro.14

So “adversary evaluation” (AE) with an integration or synthesis (AE+) could be seen as an “integrative evaluation”, emphasizing the characteristics of the phases of the evaluation process rather than a type of relation among evaluators and their arguments. The approach has a heuristic value, particularly for summative evaluations, encouraging evaluators to search for all evidence concerning the intervention that is evaluated. For the credibility of the evaluation it is important to ensure that all the evidence, as well as the different judgements, are presented in a transparent way.

In the case of formative evaluations, the proposals for improvement of the intervention would be the “thesis”, whereas the critiques of those proposals could framed as “antithesis”, and the

---

14 I am grateful to Prof. George Julnes for suggesting the link between integration and levels, and for the indication that there are different types or forms of integration, such as integration of competing causal claims, as in the first example, and integration via moderated relationships, as in the second example.
“synthesis” may include those improvement proposals that are considered of sufficient value\textsuperscript{15}, taking into account the critiques.

It should be noted that according to some definitions of AE, like Michael Scriven’s, the synthesis may or may not be included in the AE. One possibility is to involve decision makers and evaluators in a setting with clear rules of engagement to discuss alternative formulations of the synthesis. The engagement of decision makers and other stakeholders, including people who may be affected by decisions made during or after the intervention, is important to understand their perspectives. Furthermore, stakeholders’ involvement in the process of developing the synthesis is important to ensure that evaluators fully understand the environment in which recommendations will have to be implemented, and decision makers develop ownership of the evaluation’s conclusions and recommendations, thus facilitating the use of the evaluation. This involvement of decision makers does not jeopardize evaluators’ independence if the rules of engagement are clear with respect to the responsibility of the evaluators, and if they are strictly followed\textsuperscript{16}.

The following scheme shows different types of AE/Integrative Evaluations, without exhausting all possibilities:

Table. 2 Different types of AE / Integrative Evaluations

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a)</td>
<td>Two or more evaluation teams, &amp; no integration or synthesis is produced</td>
</tr>
<tr>
<td>b)</td>
<td>A single evaluation team (with sub teams) &amp; without integration or synthesis</td>
</tr>
<tr>
<td>c)</td>
<td>A single evaluation team, &amp; an integration or synthesis is produced</td>
</tr>
</tbody>
</table>

\textsuperscript{15} On “value” see Mazzucato (2018) and Moore (1995).

\textsuperscript{16} These rules could be, for example, the following: I) Only evaluators are responsible for writing the draft evaluation report; II) Decision makers involved in the implementation of the intervention that is evaluated should be interviewed and their views may be included in the report as footnotes when the evaluators disagree and III) Stakeholders’ views should be taken into account. See (Feinstein, 2019)
Traditional AE would correspond to case a), whereas case b) would also be an AE with a single evaluation team. Case c) includes a synthesis, it is an integrative evaluation (IE).

There is also the possibility of combining or blending IE with other evaluation approaches, e.g., with “realist evaluation”. The value added of IE would be to direct the evaluators’ attention not only to opposing views (as AE would do) but also to the possibility of integrating those views or different perspectives. Realist evaluations usually present “context-mechanisms-outcome” patterns without integrating them in a synthesis.

Another example of the integration of different approaches is between “goal-free evaluation” (GFE) and “objectives-based evaluation” (OBE).

The latter is the standard evaluation approach used by evaluation offices of multilateral development banks and several UN agencies. It is based on the objectives of the policy, program or project (“intervention” is the term used to refer to either policy, program or project). GFE has been presented by Scriven (1974) as a way to overcome limitations of OBE, particularly its focus on intended objectives which may leave out of focus unintended results, either positive or negative, that could be more significant than intended results.

Due to its focus on intended objectives, OBE may seem to be suitable for accountability, whereas GFE could seem more appropriate for learning (even though OBE could also be a tool for learning about intended results). Frequently GFE practitioners reject OBE, whereas those that pursue OBE generally neglect GFE.

17 See Thomas & Chindarkar (2019, chapter 4) for a presentation of OBE.
Given that GFE and OBE are compatible and complementary they can be integrated. Such an integrated evaluation approach combines the learning and the accountability functions of evaluation. The argument for the integration of GFE and OBE can be presented through the following five propositions:

1) There are Intended and unintended results

2) Objectives are a subset of intended results

3) Objectives based evaluation (OBE) focuses on a subset of intended results

4) Goal Free Evaluation (GFE) focuses on unintended results

5) OBE can be integrated with GFE

Structuring the Integrative Exchange Between Evaluators & Stakeholders

Whereas the previous section showed the importance of involving the decision makers in the process of formulating recommendations, it is worthwhile to consider how the exchange between evaluators and stakeholders can be structured so as to better examined and take into account their underlying values.

Evaluators can structure the integrative exchange between evaluators and stakeholders to discuss issues related to the implementation and results of the intervention that is evaluated using the two phases of the integrative evaluation approach. The agenda of these meetings could start with a question concerning what matters for the stakeholders, that is about their values. The answers could be taken as theses, whereas the evaluators could reverse the thesis in order to generate an antithesis. For example, if a key value elicited in the process is “equality”, the evaluators could ask about what is that the stakeholders dislike or reject about “inequality”. The discussion may yield as a synthesis “equality of opportunity”, which could
be used as a criterion to judge the value of the intervention. The same structure of the dialogue could be used to surface and triangulate values and other issues, including plausible rival hypotheses concerning the outcomes of interventions.\textsuperscript{18}

**Integrative Evaluation, Evaluative Thinking and Deliberative Democracy**

Given the importance of a discussion about values and valuing for evaluative thinking\textsuperscript{19}, and the structure that an integrative approach provides to the evaluator for a systematic search for values and insights, it can be argued that integrative evaluation facilitates evaluation thinking.

Furthermore, this way of applying evaluative thinking “can (…) help deliberators distinguish those disagreements that arise from genuinely incompatible values from those that can be more resolvable than they first appear”, which is important for a well functioning deliberative democracy.\textsuperscript{20}

For example, an integrative evaluation may show that different judgements of a universal basic income scheme applied in a given country may not be actually due to differences in values but to different judgements concerning the efficacy of this intervention.

Thus, integrative evaluation can nurture the democratic debate, strengthening deliberative democracy (Lester, 1998), which would be particularly timely given the current polarization of most societies. In contrast, “agonistic democracy”, in which opposite positions are

\textsuperscript{18} Manin (2017) and (2011/2) proposes practical ways of promoting adversarial deliberation, particularly the organization of debates disconnected from electoral competition. Evaluations are not considered by Manin but could play a key role for evidence-based deliberations, whereas integrated evaluations through their synthesis phase may facilitate the emergence of an outcome from these deliberations, integrating the adversarial positions

\textsuperscript{19} See Vo and Archibald (2018) and Schwandt (2018)

considered without attempting any integration or synthesis, leads to an intensification of conflict (Hvidsten, 2018).

So an integrative approach to evaluation can help to raise the level of deliberation concerning interventions, facilitating the convergence of different positions by replacing parallel monologues with integrated evidence-based arguments. In this way it enriches deliberative democracy through a participatory evaluation process oriented towards the production of a synthesis of adversarial positions. It is worthwhile to mention that “integrative democracy”, an emerging alternative formulation of deliberative democracy that is based on integration rather than deliberation (Wright, 2019), may be reinforced by integrative evaluation by co-creating a synthesis of different perspectives.

One of the anonymous reviewers of a previous version of this article raised an important issue: “what if ‘one side’ has facts and moral values while the ‘other side’ has ‘alternative facts’ and dubious values? Over the past four years, in the U.S., for example, the ‘both-sideism’ of the media has been critiqued for trying to frame public disputes as rational ones about policy disagreements, when, in actuality, one side of the debate sometimes explicitly espoused antisemitic, racist, xenophobic, fascistic values”.

In situations like this one it is most unlikely that there could be a dialogue and it would not make sense to attempt an integration of these perspectives as they are incompatible, irreconcilable accounts of reality. Following this line of argument, if it can be shown that the adversary evaluations (or different hypotheses, H1 and H2) are compatible and complementary then a synthesis or an integration would be feasible. Otherwise, an integrative evaluation could not be performed.

If (H1 & H2) are compatible and complementary, then IE (H1&H2) is feasible
When the views concerning facts are inconsistent, no integration is possible whereas it may be possible if the views differ on which facts are relevant; this in turn may be due either to a difference in values and/or to differences in the context. Although there is an entanglement of facts and values, as argued in (Feinstein, 2019b), for the sake of clarity in this article the entanglement is not considered.

**Comparison of Integrative Evaluation with other Approaches**

There are several evaluation approaches that are deliberate attempts to integration and therefore partially overlap with “integrative evaluation”, as shown in Table 3. The table does not include E.House’s “Deliberative Democratic Evaluation” because it is an approach that does not consider an explicit phase of synthesis, although it appropriately focuses on inclusion, dialogue and deliberation, which are consistent with “Integrative Evaluation”. Also Donald T. Campbell’s focus on “plausible rival hypothesis /or explanations/” (PRH) is not included in the table given that it does not entail a synthesis. However, when an intervention is implemented in different contexts, in each of them one of the PRH could be selected, and the synthesis would integrate those hypotheses or explanations that were considered most adequate (“a better fit”) in the different contexts.

Table 3: Comparison of Integrative Evaluation with other Approaches

<table>
<thead>
<tr>
<th>AUTHORS/APPROACH</th>
<th>SYNTHESIS OF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael Scriven/Jane Davidson</td>
<td>Dimensions/subdimensions/criteria/rubrics</td>
</tr>
<tr>
<td>Jennifer Greene</td>
<td>Methods/approaches</td>
</tr>
<tr>
<td>Integrative Evaluation</td>
<td>Complementary hypothesis/approaches</td>
</tr>
</tbody>
</table>
Conclusion

“Integrative evaluation” is an approach with two main phases: identification of plausible rival hypotheses and integration of rival hypotheses. “Adversary evaluation” (AE) may be reframed in cases in which the different perspectives are compatible and complementary as “integrative adversary evaluation”, with two phases: a first one identifying plausible rival hypotheses or approaches or perspectives, and a second phase integrating these perspectives. It can be used as a heuristic approach for summative evaluations, and also with formative and/or developmental evaluations. When the different perspectives are not adversarial but are compatible and complementary, it is possible to integrate them in a synthesis. Integrative evaluations may contribute to mitigate polarization and to enhance deliberative democracy.

Declaration of Conflicting Interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References


Feinstein, O. (2019a) Checklist for evaluation recommendations
https://wmich.edu/evaluation/checklists


Harvey, L., (2019) Social Research Glossary, Quality Research International
http://www.qualityresearchinternational.com/socialresearch/


Manin, B. (2011/2) “Comment promouvoir la délibération démocratique ?” Raisons politiques 42: 83 – 113


Ortega y Gasset, J. (1938) El tema de nuestro tiempo Madrid: Espasa Calpe


https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=2886&context=journal_articles


