

# Evaluation Criteria for Evaluating Transformation: Implications for the Coronavirus Pandemic and the Global Climate Emergency

American Journal of Evaluation  
2021, Vol. 42(1) 53-89  
© The Author(s) 2020  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/1098214020933689  
journals.sagepub.com/home/aje



Michael Quinn Patton<sup>1</sup> 

## Abstract

Fundamental systems transformations are needed to address the global emergency brought on by climate change and related global trends, including the COVID-19 pandemic, which, together, pose existential threats to the future of humanity. *Transformation* has become the clarion call on the global stage. Evaluating transformation requires criteria. The revised Organization for Economic Cooperation and Development/Development Assistance Committee criteria are adequate for business as usual summative and accountability evaluations but are inadequate for addressing major systems transformations. Six criteria for evaluating transformations are offered, discussed, and illustrated by applying them to the pandemic and the Global Alliance for the Future of Food. The suggested criteria illustrate possibilities. The criteria for judging any intervention should be developed in the context of and aligned with the purpose of a specific evaluation and information needs of primary intended users. This article concludes that *the greatest danger for evaluators in times of turbulence is not the turbulence—it is to act with yesterday's criteria.*

## Keywords

criteria, transformation, systems change, complexity, global

We are the first generation to know we are destroying our planet and the last one that can do anything about it.

Tanya Steele (2018, CEO, World Wildlife Fund)

---

<sup>1</sup> Blue Marble Evaluation, Utilization-Focused Evaluation, St. Paul, MN, USA

## Corresponding Author:

Michael Quinn Patton, Blue Marble Evaluation, Utilization-Focused Evaluation, 740 Mississippi River Blvd. South Suite 21E, St. Paul, MN 55116, USA.

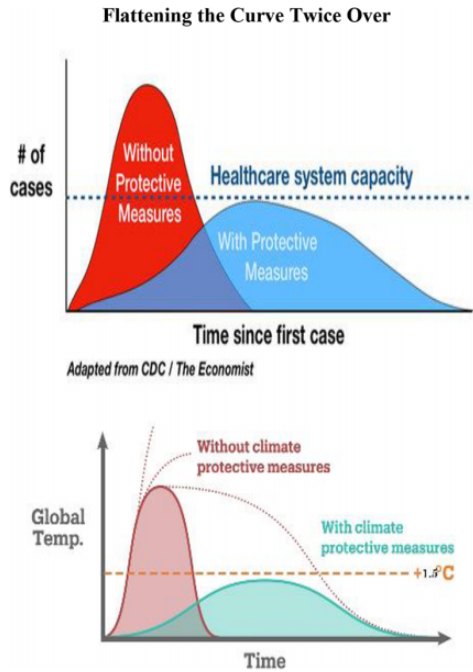
Email: [mquinnp@gmail.com](mailto:mquinnp@gmail.com)

Humanity faces a global emergency of multiple intersecting and accelerating trends. The coronavirus pandemic has provided a glimpse into the magnitude of changes set in motion by a global emergency. United Nations Secretary-General Antonio Guterres (2020), among many others, has warned consistently throughout the pandemic that climate change looms over the world as a larger, more far-reaching global emergency for which COVID-19 has been but a dress rehearsal, an early warning of what lies ahead at greater magnitude though slower manifestation. Intensifying and magnifying the global emergencies of the pandemic and climate emergency are the economic depression brought on by the pandemic lock-down and, in May/June 2020, the global uprising protesting social injustice, police violence, and systemic racism sparked by the murder of an unarmed African-American, George Floyd, by a white police officer in Minneapolis. Accompanying these intersecting and mutually reinforcing global change processes is a dramatically escalating *infodemic* of misinformation, fake news, distortions of facts, outright lies by some policy makers, attacks on and disregard for science, and weakening institutional accountability manifested most visibly by the Trump administration firing of senior inspectors general. The pandemic offers but one window into the implications for evaluation of these global processes.

Evaluation responses to the pandemic were widespread and immediate but largely ad hoc and reactive (Independent Evaluation Group [IEG], 2020b; Patton, 2020b; Tolley, 2020). Chelsky and Kelly (2020) of the World Bank described monitoring and evaluation (M&E) during the pandemic as “bowling in the dark.” Better Evaluation (2020) offered systematic and comprehensive guidance for adapting evaluation’s response to COVID-19 based on the dimensions of the Rainbow Framework for Evaluation. All evaluation association conferences planned for 2020 had to be altered or canceled, and all associations issued statements about the continuing importance of evaluation and support for evaluators. Much evaluation training was moved online. Evaluators have been reflecting and blogging furiously and thoughtfully about what the pandemic and the climate emergency mean for evaluation (e.g., Bitar, 2020; Chaplowe, 2020; Feinstein, 2019, 2020; IEG, 2020a; Ofir, 2020; Ramalingam et al., 2020; Vidueira, 2020). Efforts abound at drawing lessons from the pandemic to inform the response to climate change (e.g., Cartier, 2020; Euber, 2020; Karalisi, 2020).

It is also instructive, and I must say disheartening, to look backward at what has been learned—and what has since been ignored—about evaluating prevention of HIV and mitigating that global epidemic (Rugg et al., 1999), lessons that have taken on added relevance and significance in addressing the COVID-19 crisis despite the different nature of HIV transmission. Fundamental prevention and mitigation principles flowing from epidemiology and evaluation still apply, ignored though they may be by contemporary politicians (Mukherjee, 2020). For example, the Centers for Disease Control and Prevention (CDC) *Field Epidemiology Manual*, developed scientifically over the course of decades, provides detailed protocols for dealing with all aspects of a pandemic, including communications with the public. That knowledge and wisdom were largely ignored by politicians in the United States as the CDC was muzzled throughout the pandemic emergency (Duhigg, 2020).

The global pandemic has provided substantial evidence to reinforce and make even more urgent the premise that major systems transformations are needed to address the global emergency brought on by climate change and related global trends. Global warming; pollution of oceans, land, and air; biodiversity loss; species extinction; and virulent infectious diseases pose existential threats to the future of humanity (Kolbert, 2020; United Nations Environment Program, 2019). *The Economist* featured a cartoon showing two boxers fighting, one with the head of the world and the other with the head of the coronavirus. Observing the fight from outside the ring, but looming menacingly over it, was a much larger boxer with a fiery head wearing trunks labeled “climate change.” The widely communicated and highly effective graphic created by CDC (2020) depicting the need to “flatten the



**Exhibit 1.** Flattening the curve twice over. Source: Hayes (2020), NYU Energy & Environmental Impact Center (<https://www.weforum.org/agenda/2020/04/flattening-the-climate-curve-in-the-postcovid-world/>).

curve” to fight coronavirus has been redrawn to communicate the urgent need to flatten the curve of global warming (see Exhibit 1).

The pandemic has been global in scale and universal in impact. The climate emergency will, likewise, be global in scale and universal in impact. The global climate emergency affects all of us, leading to calls for action in whatever niche we inhabit. For evaluators, that niche is evaluation. Evaluation has emerged as critically important in realizing the vision and aspirations of the sustainable development goals (SDGs; Rugg, 2015, 2016; Eval4Action, 2020, Patton, 2020c). But current evaluation criteria and practices are inadequate for evaluating transformation. That makes evaluation part of the problem. Evaluating transformation means transforming evaluation. Toward that vision, I’ll offer six criteria to guide evaluating transformation.

## Transformation

We stand on the brink of a technological revolution that will fundamentally alter the way we live, work, and relate to one another. In its scale, scope, and complexity, the transformation will be unlike anything humankind has experienced before. (Schwab, 2017, *World Economic Forum*)

*Transformation* has become the clarion call on the global stage. Humans are using the Earth’s resources at levels, scales, and speed that are changing Earth’s ecological systems and, in so doing, warming, polluting, and degrading the environment at a level that threatens the future survival of humanity. Charly Kleissner (2020), an “impact entrepreneur” and a leader in the impact entrepreneur network provides a perspective from the financial community:

We are at the beginning of an unprecedented global transformation, where humanity has precious little time to figure out how to sustainably live on a planet with finite resources. During this transition to an ‘impact economy’ all systems will have to change, especially the financial system. The status quo of adapting existing systems to deal with systemic issues like inequality, social justice, climate change and poverty is not going to be good enough.

One could fill this journal with such quotes. A substantial literature documents the nature and extent of the global climate emergency and the need for and nature of the transformations required (e.g., Fazey, 2020; Fazey et al., 2018; Mora et al., 2018; Patton, 2019a, 2019c, 2020a). The world faces a global emergency from the combination of our warming climate; virulent infectious diseases; pollution of land, air, and water; dying coral reefs; millions of displaced people; rising inequalities; cyberterrorism; pandemic threats; increased nuclear dangers; ever more severe weather; species extinction; biodiversity loss; white supremacy; and nationalist ideologies—and related trends and challenges.

### **Evaluation and Transformation**

The field of evaluation has not been oblivious to climate change ( Picciotto, 2009; van den Berg & Feinstein, 2019, 2020) nor to the global call for transformation (Picciotto, 2020a). The IEG (2016) has generated a framework for “supporting transformational change for poverty reduction and shared prosperity.” The Global Environment Facility (2018; Independent Evaluation Group, 2018) has developed innovative evaluation methods to address climate change and other environmental issues. The theme of the 2014 Annual Conference of the American Evaluation Association (AEA) was *visionary evaluation for a sustainable, equitable future*, now a book (Parsons et al., 2020). A special issue of *New Directions for Evaluation* was devoted to *evaluating sustainability* (Julnes, 2019). The 2018 theme of the Australasian Evaluation Society was *transformations* while that of the European Evaluation Society was *evaluation for more resilient societies*. The theme of the 2019 Conference of the International Development Evaluation Association (IDEAS) was *evaluation for transformative change*, supported by a publication with that title (van den Berg et al., 2019). *Blue Marble Evaluation* for global systems transformations posits that evaluating transformation requires transforming evaluation (Patton, 2020a). Evaluating transformation is now firmly inscribed on the global evaluation agenda.

### **Evaluation Criteria**

Before discussing specific criteria for evaluating transformation, let me remind readers of the centrality of criteria to evaluation. We can’t evaluate without criteria. Indeed, we can’t communicate with each other without criteria. How do you want your breakfast eggs? The answer requires criteria. How will you judge the value of this article? What kind of evaluator are you? Criteria matter.

1. Criteria constitute the nuclear core of evaluation’s energy function: rendering judgment. Without criteria, there can be no judgment. Without judgment, there can be no evaluation.
2. Criteria express, manifest, encompass, make explicit, and operationalize what is valued. Criteria mediate the conversion of values into judgments.
3. Criteria prioritize what is important.
4. Criteria direct what questions to ask, data to collect, and results to highlight.
5. Criteria focus evaluation reporting and conclusions: To what extent and in what ways have criteria been met?

## Criteria and the Logic of Evaluation

The formal logic of evaluation involves four steps:

1. define the criteria that will be used to evaluate something;
2. set standards of performance on those criteria;
3. measure the actual performance; and
4. synthesize the results to reach an evaluative judgment (Azzam, 2018; Davidson, 2005; Scriven, 2016).

How deeply embedded is this logic among evaluators? To find out, Ozeki et al. (2019) surveyed members of the AEA in two separate random samples. While most respondents were not familiar with the four-step logic as a formal evaluation framework, most reported following the steps in their practice: 82% reported identifying evaluative criteria.

The logic of evaluation applies to evaluating interventions (projects, programs, policies, and initiatives) and to evaluating evaluations (meta-evaluation). The Joint Committee Standards for Educational Evaluation (2010) specify that evaluations should be judged by their utility, feasibility, propriety, accuracy, and accountability. The AEA Guiding Principles (2018) call for evaluations to be judged by the systematic nature of the inquiry, evaluator competence, integrity, respect for people, and contribution to an equitable and just society. The evaluation standards and principles specify criteria for judging the quality of evaluations.

Criteria for judging interventions specify how to judge the merit, worth, and significance of whatever is being evaluated, what we call the *evaluand*. The most influential and widely used criteria for evaluating development interventions are those adopted and disseminated by the Development Assistance Committee (DAC) of the Organization for Economic Cooperation and Development (OECD). As it happens, just as the extent and implications of the global climate emergency have become more evident in the last couple of years, the DAC criteria were undergoing revision. Reviewing the DAC criteria will establish the context for considering different and more focused criteria for evaluating transformation.

## DAC Evaluation Criteria

The OECD DAC, based on work begun in 1984, developed, endorsed, and disseminated five evaluation criteria in 1991: relevance, effectiveness, efficiency, impact, and, sustainability. These criteria became widely adopted as defining evaluation quality for evaluating all international development and humanitarian projects, programs, and policies (Kennedy-Chouane, 2020; Picciotto, 2013). Niels Dabelstein, a distinguished thought leader in international development evaluation, wrote a detailed history of 30 years of the work of the DAC Network on Development Evaluation (OECD, 2013). He concluded the DAC criteria have “had a profound impact on development evaluation. These have become the widely accepted criteria upon which every development evaluation will base its assessments—or make excuses for why not” (OECD, 2013, p. 33).

Robert Picciotto, a former director of the World Bank’s IEG and a founder of both the International Program for Development Evaluation Training (IPDET) and the IDEAS, believes that the DAC criteria contributed significantly to improving the quality of development evaluations, especially in formulating multiple criteria that went beyond the singularly dominant criteria of goal attainment and return on investment. “Historically, the addition of relevance, sustainability and impact to the list made a very big difference. It was nothing short of revolutionary” (Picciotto, 2020b). He goes on to caution that “criteria are a floor rather than a ceiling. The DAC criteria are necessary but far from sufficient.” Criteria are not enough to guarantee quality. As Megan Kennedy-Chouane (2020), the Head of Evaluation, Development Co-operation Directorate, Review, Results,

Evaluation & Development Innovation, OECD, put it in a webinar introducing the revised DAC criteria: “Good evaluations can be conducted with bad criteria and bad evaluations can be conducted with good criteria.” Still, the criteria matter. The quality of the “floor” affects the solidity of the above-floor structure.

### *Revising the DAC Criteria*

In light of the Global Agenda for 2030 and the SDGs, DAC began reviewing whether the criteria needed revision. A consultation process from March to October 2018 included interviews with key stakeholders, a consultation workshop, discussions at international meetings/seminars throughout the world, an OECD DAC Network member survey, and a public survey with stakeholders. The survey received 691 survey responses with over 700 pages of qualitative comments. In addition, 11 development agencies submitted formal commentaries. The concluding survey item asked whether the current DAC criteria should be retained, adapted, or removed. Over 89% recommended retaining the existing criteria (DAC Network on Development Evaluation, 2018).

Building on this feedback, the revised criteria were published in late 2019 in a report entitled *Better Criteria for Better Evaluation*. The document lays out adapted definitions for relevance, effectiveness, efficiency, impact, and sustainability—and for one new criterion, coherence (OECD/DAC, 2019a). The document makes clear that the criteria are intended to be universally applied to all interventions:

The criteria play a normative role. Together they describe the desired attributes of interventions: all interventions should be relevant to the context, coherent with other interventions, achieve their objectives, deliver results in an efficient way, and have positive impacts that last. (OECD/DAC, 2019a, p. 5)

#### **Exhibit 2.** Revised DAC Criteria.

---

*Relevance: Is the intervention doing the right things?*

---

The extent to which the intervention objectives and design respond to beneficiaries’, global, country, and partner/institution needs, policies, and priorities and continue to do so if circumstances change.

---

*Coherence: How well does the intervention fit?*

---

The compatibility of the intervention with other interventions in a country, sector, or institution.

---

*Effectiveness: Is the intervention achieving its objectives?*

---

The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups.

---

*Efficiency: How well are resources being used?*

---

The extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way.

---

*Impact: What difference does the intervention make?*

---

The extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects.

---

*Sustainability: Will the benefits last?*

---

The extent to which the net benefits of the intervention continue, or are likely to continue.

---

*Source.* Organization for Economic Cooperation and Development/Development Assistance Committee Network on Development Evaluation (2019a).

Exhibit 2 presents the six revised DAC evaluation criteria. Since they have been published, the revised criteria have been widely disseminated and supported with webinars and technical assistance (e.g., Kennedy-Chouane, 2020).

### *Critique of the DAC Criteria: A Business as Usual Message*

The revised criteria are useful for those who want to continue designing and evaluating projects and programs in familiar, comfortable, well-known, and well-traveled ways. The revision amounts to some fine-tuning and tweaking but is basically business as usual. What was once innovative and even “revolutionary” has become staid, conservative, and common place. The criteria apply primarily to summative and accountability purposes of evaluation, important to be sure, but hardly relevant to the great panorama of evaluation approaches and purposes that have emerged in the last three decades. The unchanged labels of the five DAC criteria—relevance, effectiveness, efficiency, impact, and sustainability—carry the message that things can go on as before. You have to study the revised definitions carefully and scrutinize them diligently to detect any meaningful substantive changes. I suspect that few who use the current criteria, whether in commissioning evaluations, conducting them, or judging their quality (meta-evaluation) will notice any differences or study the accompanying guidance document meant to explicate adaptations. The addition of “coherence” as a new criterion addresses some of the limitations of the original five, meant to address climate change, poverty reduction, equity, and other goals at whatever level the evaluation is occurring, based on the goals of the intervention (Kennedy-Chouane, 2020). But “coherence” doesn’t evoke or connote, at least for me, the nature of complexity or the magnitude and urgency of transformation. You have to dig deep and interpret broadly to arrive at that understanding. The very label “coherence” evokes orderly stability. Indeed, the overall message of the DAC criteria revision is *business as usual*. Carry on. Continuity reigns supreme.

Embracing continuity is a common phenomenon among longtime users of systems and frameworks of all kinds. Ascendance leads to dominance leads to inertia. Writing about creating more resilient and adaptable systems, former Rockefeller Foundation President Judith Rodin (2014) observes: “Systems want to remain stable, to continue on as they always have” (p. 52). In a similar vein, Clayton Christensen, the pioneer of the theory of disruptive innovation, has posited that successful enterprises have a strong vested interest in serving long-term customers and supporters. Incumbent leaders of successful efforts and organizations listen to their existing customers and concentrate on sustaining their past success, thereby resisting innovation (Christensen et al., 2015). Applied to DAC, the widespread adoption of the DAC criteria, including being widely mandated for use in development evaluations worldwide (Kennedy-Chouane, 2020), created a build-in conservative inclination to resist innovation. The consultation responses resisted change. Tweaking, yes. Major reform, no. Thus, the revised DAC criteria remain mired in the logic of “normal evaluation” fixated on effectiveness, efficiency, control, and continuity (Schwandt, 2019; Vataja & Parkkonen, 2019). In contrast, “postnormal” evaluation acknowledges, confronts, and engages with complexity, unpredictability, wicked questions, incompleteness, contradictions, turbulence, instability, “a plurality of perspectives in value determination” (Schwandt, 2019, p. 317), systems thinking, and, quintessentially, transformation. Normal evaluation relishes the lessons that flow from hindsight; however, alluring, and seductive but ultimately ephemeral, such lessons are generalizations that decay more rapidly than the half-life of radioactive particles. Evaluation for transformation requires expanding *foresight evaluation capacity* grounded in the ethical responsibility of having *skin-in-the-game* (Patton, 2019b, 2020; Taleb, 2018) because we are all affected by how, and how well, humanity endures.

It is worth noting that the survey soliciting feedback about whether the DAC criteria needed revision provided no context by way of mentioning the global climate emergency or calls for transformation. Might there have been more support for significant reform of the DAC criteria if

a transformational context had been provided? Might reactions have been different if the consultation process had occurred during the pandemic? We'll never know. Such a context was not provided, so what emerged was reaffirmation of normal evaluation in support of the status quo and a modest openness to marginal, incremental change.

While I am critical of the failure of the revision process to identify criteria that are more directly and boldly relevant to our postnormal world, and the corresponding need for postnormal evaluation, it is important to recognize and applaud the broader efforts of DAC to support high-quality evaluation. For example, DAC has long advocated and supported strengthening the evaluation culture of organizations. DAC has promoted 12 lessons in that regard that remain relevant, important, and insightful.

1. Base development policy decisions on evidence.
2. Make learning part of the culture of development co-operation.
3. Define a clear role for evaluation.
4. Match ambitions with adequate resources.
5. Strengthen program design and management systems.
6. Ask the right questions and be realistic about expected results.
7. Choose the right evaluation tools.
8. Work together.
9. Help strengthen partner country capacities and use them.
10. Act on evaluation findings.
11. Communicate evaluation results effectively.
12. Evaluate the evaluators (Lundgren, 2019, p. 13).

### *Alternative Pathways Forward*

The issue before us is criteria for evaluating transformation. It seems to me that there are two options. One approach would be to interpret the revised DAC criteria in ways that make them relevant to transformation. That is what the guidance document recommends by including transformation under the impact criterion.

Impact addresses the ultimate significance and potentially transformative effects of the intervention. It seeks to identify social, environmental and economic effects of the intervention that are longer term or broader in scope than those already captured under the effectiveness criterion. (OECD/DAC Network on Development Evaluation, 2019b, p. 11)

However, to the credit of DAC dissemination effort, the OECD website makes it clear that the revised criteria are less than optimal for addressing transformation. In responding to “frequently asked questions,” the issue is addressed directly:

Can these criteria be used to evaluate transformational change?

The DAC reply:

Yes and no. Yes, the criteria can be used to understand how different interventions are (or are not) creating transformational changes. The potential transformative effects of an intervention are captured under the new definition of impact. We expect evaluations will look more and more at questions related to transformation, reflecting the growing interest in understanding transformational change and the imperative for transformation described in both Agenda 2030 and the Paris Agreement. The evaluation methods needed to answer these questions are still emerging.



However, the evaluation criteria are intended to be used in evaluating an intervention (some intentional effort to create a change in the world), they are not a particularly useful tool for descriptive analysis of transformative change or systems change. (OECD, 2020)

The alternative to forcing the new wine of transformation into the old bottles of the DAC criteria is to elevate attention to transformation by developing criteria that specifically highlight the nature, scope, and breadth of changes connoted by the term *transformation*. Responding to the systemic threats of the pandemic and climate emergency requires audacity: emergency responses, by definition, disrupt business as usual mindsets, modalities, and methods. Yet, policy makers have yet to grasp the nettle, and evaluators had been mostly going about their evaluations in a business as usual mode, at least until the pandemic ended the pretense that “normal” was a viable future and pushed the whole world into uncertainty about what the future holds. We now live and work in a *business as unusual world*, a postnormal world, a global emergency world, a time-is-running-out world. In what follows, I offer examples of alternative criteria to suggest what transformation-specific criteria might constitute and communicate. The criteria offered here result from 2 years of reflection, consultation, workshopping, and feedback about criteria for transformation with others.<sup>1</sup> In sharing them here, I mean for them to illustrate possibilities and stimulate further contextual adaptation, not to be treated as universal, standardized, and/or mandated criteria.

To demonstrate the applicability of these criteria, I will briefly comment on how they could have been applied in evaluating pandemic interventions. I’ll also offer concrete illustrations of how they can, and are, being used by the Global Alliance for the Future of Food (Global Alliance as shorthand). The Global Alliance’s 30-member philanthropic foundations collaborate to support transformation of food and agricultural systems. Working with the Global Alliance to support strategic evaluation has influenced the development of these criteria. Thus, following presentation and explanation of each criterion, I’ll show how it applies to the Global Alliance.

## Evaluation Criteria for Evaluating Transformation

### *Transformation Fidelity Criterion*

The extent to which the realities of transformational change initiatives match transformational aspirations and rhetoric.

#### *Evaluation implications.*

Ensure that what is called transformation constitutes transformation.

*Evaluate whether and how what is called transformational engagement constitutes a trajectory toward transformation.*

#### *Explanation and elaboration.*

Those development programs that are most precisely and easily measured are the least transformational, and those programs that are most transformational are the least measured.

Andrew Natsios (2010, p. 1), Administrator [Agency Head]

U.S. Agency for International Development, 2001–2006 (p. 1)

Evaluation has a long history with the criterion of *fidelity*. Fidelity evaluation includes assessing integrity and rigor in replicating effective programs to new localities (Are the replications faithful to the original model on which they are based?) Just as fidelity is a central issue

in efforts to replicate program models, evaluation fidelity concerns whether an evaluator following a particular model is faithful in implementing all the steps and processes of that model (Miller & Campbell, 2006). I examined fidelity to developmental evaluation (DE) principles that opened with an experienced DE practitioner telling me: “More often than not, I find, people say they are doing Developmental Evaluation, but they are not” (Patton, 2016).

There is a lot of hype around transformation, as the term has become widely used and taken on a trendy cachet. Claims of transformation abound. Ensuring that such claims are meaningful and consistent with the face validity of the construct becomes a transformational evaluation priority under this criterion. Thus, the fidelity criterion aims to bring some rigor to the very notion of transformation.

The IEO of the Global Climate Facility has defined “transformational change as deep, systemic, and sustainable change with large-scale impact in an area of global environmental concern.” Operations that are “merely” highly successful, complex, or large in size are not, by definition, transformational. To be transformational the intervention (1) must address a global environmental challenge such as climate change, biodiversity loss, or land degradation; (2) bring about fundamental and deep change in a system or market; and (3) create large-scale change that (4) is financially, economically, socially, and politically sustainable in the long term after the intervention ends (IEO, 2018).

The core evaluation question for the fidelity transformation criterion becomes: *What is being envisioned as transformational and how is that vision being realized through action?* This question could be addressed under the DAC criteria of relevance or impact, but the point of articulating transformation-specific criteria is to call attention to the scale, scope, urgency, and challenges of the global emergencies. Relevance and impact are generic and might, or might not, include transformation. Transformation fidelity puts transformation front-and-center in the context of the realities of global emergency. Experienced development evaluator Osvaldo Feinstein (2019) has emphasized this point writing about evaluation for transformational change:

An intervention (say a policy) may be highly relevant and get high marks on the other four [DAC] criteria, and yet it may not change in any significant way the situation of the target population. Its actual effects, although positive, may be insignificant in terms of transformational change. (p. 19).

I treat transformation as a sensitizing concept. It can not be operationally defined and measured in a standardized way applicable to any context. Rather, it is a concept that has to be given meaning and specificity within the context where transformation is targeted. The transformation fidelity criterion requires a contextually appropriate definition for any initiative claiming a transformational purpose.

*Is an initiative, or more likely a set of initiatives and interventions, on a trajectory toward transformation?* Asking the trajectory question changes the evaluation focus from transformation having occurred (or not), as the DAC criteria for effectiveness and impact specify, to *transformational engagement*. That is the reframing formulated by the influential IEG of the World Bank. Assessing the trajectory toward transformation is what most funders, decision makers, and implementers of initiatives are looking for from evaluation.

Transformational engagement is an intervention or a series of interventions that helps achieve deep, systemic, and sustainable change with large-scale impact in an area of a major development challenge. These engagements help clients remove critical constraints to development; cause or support fundamental change in a system; have large-scale national or global impact; and are economically, financially, and environmentally sustainable. (IEG, 2016, p. 1)

The IEG of the World Bank evaluated a sample of 20 transformational engagements varying in form, size, the development challenges they address, sector, and region as well as country

context. In addition, IEG reviewed a purposeful and selective sample of country-level engagements. Their comparative and synthesis analysis exemplifies applying the transformation fidelity criterion (Independent Evaluation Group, 2016; Independent Evaluation group, 2017; see also Heider, 2017).

The transformation fidelity criterion includes examining whether the purported transformational engagement is based on an evidence-supported theory of transformation (Patton, 2020a, chapter 13). It is beyond the scope of this article to present a theory of transformation. I would simply suggest at this point that **a theory of transformation synthesizes multiple theories of change.** Any specific theory of change concerns how to produce desired results targeted by a particular intervention. **Transforming systems requires aligning, networking, and integrating multiple and diverse theories of change to build critical mass transformational tipping points.** Transformation, then, is not an intervention, it is rather a movement creating synergies among multiple interventions (Patton, 2020a). That brief vision of a theory of transformation points toward what transformational fidelity means. Transformation is not a project. It is major systems change.

In her Foreword to *Evaluation for Transformational Change* (van den Berg et al., 2019), Hernandez-Licona (2019) opens by positing that “the whole development system (or the multiple development systems)” must be transformed but goes on to note that the world is not on track for achieving most of the 169 targets for the SDGs.

While some goals are not on track, others are even going backwards, to rising inequalities, climate change, biodiversity loss, increasing waste from human activity, violence conflicts and related humanitarian crises causing the displacement of millions of people. We need to make more changes in the way we arrange our economies, our societies and our politics if you really want to have a better world in 2030. Business as usual will not do the trick. (p. x).

**A shorthand way of thinking about the transformation fidelity criterion is that it is about examining whether those making transformation claims are actually walking the talk.** Greenwashing (derived from whitewashing) is the phenomenon of corporations and other organizations claiming to be taking major environmentally friendly actions that are only modestly implemented, or are actually as much or more motivated by cost savings (hotels saving on laundry costs when guests reuse towels) and advertising hype (attracting consumers who are attracted to environmentally friendly products).

In essence, the transformation fidelity criterion epitomizes evaluation’s dual reality-testing and evaluative thinking mandates. **To what extent does the reality of transformational engagement match the rhetoric? To what extent and in what ways does the conceptualization of transformation reflect rigorous evaluative thinking?** The role of evaluation in asking and answering these kinds of questions is critical to a future dependent on transformation occurring. Taleb (2018), author of best-selling book *Skin in the Game*, posits that “It is easier to macrobullshit than to microbullshit.” One job of evaluators is to detect and expose the bullshit, micro, or macro. But there is so much, it is hard to keep up, as it is piling deeper and higher. **The transformation fidelity criterion makes evaluators macrobullshit detectors with regard to transformation claims.**

*Transformational fidelity and the coronavirus pandemic.* A great deal of time was lost in January and February, 2020, as many public officials, especially President Trump, downplayed the scope, scale, and significance of the pandemic. Previous infectious disease threats like SARS, MERS, H1N1, and Ebola had been contained. The ultimate long-term effects of the pandemic and its transformative dimensions are still unfolding, but as I write this in May 2020, there’s a growing consensus that there will be no return to “normal.” COVID-19 is proving transformative even though much of the response to the pandemic attempted to contain its systems-altering

significance. A major evaluation challenge looking ahead will be to track, document, and extract lessons from just how transformative the coronavirus turns out to be. Transformational fidelity as a criterion directs us to ask that question.

A team of internationally recognized experts, including Nobel Prize winner Joseph Stiglitz and well-known climate economist Nicholas Stern, came together to assess the economic and climate impact of taking a green route out of the pandemic crisis. They catalogued more than 700 stimulus policies into 25 broad groups and conducted a global survey of 231 experts from 53 countries, including senior officials from finance ministries and central banks. **Their analysis of whether COVID-19 fiscal recovery packages will accelerate or retard progress on climate change portrays the interconnection between the coronavirus pandemic, economic policies, social justice, global inequities, and environmental consequences which, taken together, portray the transformations necessary to attain a more sustainable and equitable future (Hepburn et al., 2020).**

*Global Alliance for the Future of Food application.* The Global Alliance has adopted a strategy aimed at stimulating local and global action and interaction for transformational change in collaboration with other committed stakeholders. “Transformational means realizing healthy, equitable, renewable, resilient, and culturally diverse food systems shared by people, communities, and their institutions” (Richardson, 2018). In January 2020, the Global Alliance (2020; see Exhibit 3) formally adopted a theory of transformation that informs its activities and provides a basis for evaluating its products, activities, and impacts through the lens of transformational engagement. Reviewing Global Alliance documents, reports, convenings, meetings, activities, initiatives, decisions, and evaluation approach make it clear that transformational engagement is not just rhetoric but is deeply embedded as the focus of ongoing work.

**Exhibit 3.** Global Alliance for the Future of Food.

---

Theory of Transformation Summary

---

**Genuine food system transformation takes place when diverse actions, networks, and individuals intersect across sector and issue silos, the global and local, the macro and the micro. These intersections facilitate convergence around shared visions and values and, ultimately, build critical mass and momentum behind tipping points that lead to healthy, equitable, renewable, resilient, and culturally diverse food systems that dynamically endure over time.**

---

**Complex Systems Framing Criterion**

Assess systems transformation using systems thinking principles and complexity concepts.

*Evaluation implications.*

Ensure that transforming systems is the transformational focus.

*Apply complex systems understandings, concepts, and frameworks in evaluating transformation.*

*Explanation and elaboration.*

Transformation emerges as a natural process in which our deeper intelligence can handle the inter-connected systemic nature of the world, can anticipate in the present moment patterns of the future and can, through sensitivity to context and with committed action, select from those future potentials actual

steps which are creating a better world. Systems thinking for a turbulent world is natural systems thinking. The catch is that, starting from where we are, we have to work at learning new patterns of perception, new tools of thinking, new methods of collaboration and new states of mind.

... Anthony Hodgson (2019)

Author of *Systems Thinking for a Turbulent World* (2020)

Transformation is not a project or program. Transformational initiatives are not targeted to achieving SMART goals, which is the traditional criterion of effectiveness. Transformation means changing systems, which means dealing with complexity dynamics in a world characterized by turbulence, uncertainty, unpredictability, and uncontrollability. **The focus of evaluation, the *evaluated* in our jargon, is *transformed systems*.**

Zenda Ofir has posited “the imperative of viewing development from a complex adaptive systems perspective—in my view, imperative if we want to optimise the chance of development success based on good design, implementation and evaluation” (Ofir, 2018a). This means **dealing with nonlinearities, interconnectedness, emergence, coevolution, adaptation, and path dependence** (Ofir, 2018a; see also Capra, 2004; Patton, 2011, 2015, 2020). The complex systems criterion is consistent with “systems thinking for a turbulent world” (Hodgson, 2020), **“evaluation in turbulent times”** (Furubo et al., 2013), “dealing with complexity in development evaluation” (Bamberger et al., 2016), “complexity-responsive evaluation” (Bamberger & Mabry, 2020, pp. 331–351), evaluating the complex (Forss et al., 2011), and Blue Marble Evaluation for global systems transformation (Patton, 2020a), all of which propose designing, engaging, and evaluating transformational initiatives through the complex dynamic systems lens. “A complex systems perspective broadens the parameters of ‘relevant’ evidence and theory for intervention development” (Moore et al., 2019, p. 24). **Orchestrating transformation** as a “theory for intervention development” means changing complex systems.

We are seeing emphasis on systems change wherever serious actors are addressing the climate emergency and global inequities. For example, the global financial investment community has been highlighting *assessing system-level investments* (Lydenberg & Burckart, 2020). As that report shows, trillions of dollars are being directed at systems-level change and social impact investors are seeking new approaches to assess such changes. The COVID-19 pandemic has increased the flow of private-sector funds into systems transformations (The Investment Integration Project [TIIP], 2020).

Complex systems thinking is so critical to evaluating transformation that it should not be buried or subsumed under a broad criterion like DAC’s coherence criterion. In part, it needs to be explicitly called out and highlighted because complex systems thinking poses its own evaluation fidelity problem. The Systems in Evaluation (2018) AEA Topical Interest Group spent 2 years identifying **the principles that constitute systems thinking: focusing on interrelationships, perspectives, boundaries, and dynamics. Incorporating complexity understandings does not mean ignoring linear causality, intended goals-based effectiveness, and broader community impacts, but it means seriously, systematically, and rigorously attending to nonlinearities, emergence, adaptation, dynamic interdependencies, and shifting interconnections in the context of cultural patterns, power imbalances, inequities, and forces that both move toward transformation and opposing forces that undermine transformation.** Attention to transformation means elevating complex systems thinking as essential, not optional. An example of such elevation was the 2020 Conference of the European Evaluation Society in Copenhagen, now postponed to 2021, which spotlighted complexity in its theme: *Evaluation in an uncertain world: Complexity, legitimacy, and ethics*. The first conference track is: “The Anthropocene and its complex problems: The role of evaluation.”

Osvaldo Feinstein (2019), writing about the need for “dynamic evaluation” in *Evaluation for Transformational Change* (van den Berg et al., 2019), considers complexity as part of the context for

transformation. “Structurally linear tools, such as log frames, should be replaced by an evaluation framework that takes complexity into account” (p. 25).

Some analysts treat systems thinking and complexity theory as separate and distinct frameworks. In formulating principles for DE, I identified a principle for incorporating a complexity perspective and a separate principle for systems thinking. Whether to treat systems thinking and complexity framing as separate criteria, or combine them as I have done in this instance, is part of the decision to be made for any given design and evaluation of transformation depending on what will be most useful in conceptualizing the nature of the transformational engagement and aligned evaluation.

In either case, transformation means systems change and that means that *complexity rules*. In complex dynamic systems, no one is in charge, no one can exercise control, the path forward is uncertain and unpredictable, and will surely require multiple networked initiatives and interventions and flow from some combination of both what is intended and what is emergent, both linear interventions and nonlinear reactions and interactions.

*Complex systems framing and the coronavirus pandemic.* Among many other things, the global pandemic powerfully demonstrated the interconnections among health systems, school systems, community systems, economic and finance systems, entertainment systems, and political systems. At any given moment, the focus tended to be on some discrete and particular solution like wearing masks, social distancing, more testing, quarantining the sick, and flattening the curve. But the entire health system was in crisis, an emergency that emerged and rapidly accelerated from years of neglect, ignored warnings, and underresourced existing health systems at all levels. A major debate, still ongoing, is whether the problem will be solved with a vaccine or will require major health systems transformations to prepare for future pandemics and related global climate emergency trends. The transformation fidelity criterion directs us to examine whether the actions proposed and implemented, like giving people in the United States \$1,200, actually transform systems or merely treats symptoms.

The pandemic epitomizes what it means to operate scientifically and evaluatively in a complex dynamic systems emergency. Consider the nature of epidemiology and what evaluators can learn from that esteemed and crucial profession.

Epidemiology is a science of possibilities and persuasion, not of certainties or hard proof. “Being approximately right most of the time is better than being precisely right occasionally,” the Scottish epidemiologist John Cowden wrote, in 2010.

You can only be sure when to act in retrospect . . .

Epidemiologists must persuade people to upend their lives—to forgo travel and socializing, to submit themselves to blood draws and immunization shots—even when there’s scant evidence that they’re directly at risk . . .

Epidemiologists also must learn how to maintain their persuasiveness even as their advice shifts. The recommendations that public-health professionals make at the beginning of an emergency—there’s no need to wear masks; children can’t become seriously ill—often change as hypotheses are disproved, new experiments occur, and a virus mutates. (Duhigg, 2020, p. 17)

Evaluators have much to learn from epidemiologists about how to engage in complex dynamic systems during emergencies, which is the world we likely all face with the exacerbating global climate emergency going forward even as that challenge intersects with economic recession and the social justice uprising.

*Global Alliance for the Future of Food application.* The Global Alliance strives to forge new insights and strengthen evidence for global systems change in agriculture and food. Transformation includes

changing inputs and outputs, transforming production, distribution, and consumption and changing both what is produced and how it is produced. The Global Alliance has adopted an agroecological perspective that is based on complex systems understandings and dynamics. The evaluation approach includes mapping food systems landscapes and creating and tracking social systems maps of networks and initiatives working on agriculture and food systems transformation as well as the institutions, corporations, and initiatives that oppose transformation. Ecologically based complex systems thinking is deeply embedded in the initiatives of the Global Alliance and, therefore, in its evaluation approach. This is the value of making a criterion for evaluation *complex systems framing*.

That brings us to the third criterion: eco-efficiency.

### **Eco-Efficient Full-Cost Accounting Criterion**

Document and assess the full costs and benefits of systems transformations, including economic, social, and environmental dimensions.

*Evaluation implications. Compare the full costs and benefits of baseline versus transformed systems.*

*Evaluate whether, how, and to what extent transformational engagement generates net eco-efficient benefits.*

*Explanation and elaboration.* Eco-efficiency offers a framework for examining transformation from unsustainable development to sustainability. This means looking beyond the traditional DAC efficiency criterion of examining the comparative costs (inputs) and benefits (outcomes) of an intervention within the boundaries of the intervention, essentially a closed system analysis. Eco-efficiency opens and expands the analysis to examine the effects of creating goods and offering services on the use of environmental resources, effects on ecosystems, and possible contributions to climate change, waste, a pollution, and reduced disparities. It combines attention to both environmental and economic efficiencies (Ehrenfeld, 2005; OECD Secretariat, 2002). Rowe (2019) has sounded a “call to action” aimed at evaluators to incorporate a two-system framework connecting human and natural systems to design “sustainability-ready evaluations.”

A related framework is the *Triple Bottom Line* which examines together economic, social, and environmental costs and benefits (Elkington, 2018; Slaper & Hall, 2011). The Stockholm Resilience Center has created a wedding cake representation of the SDGs in three-layered groupings: biophysical goals, social goals, and economic goals (Rockström & Sukhdev, 2016). Uitto (2019) has conceptualized an evaluation framework for addressing the SDGs by integrating three dimensions of sustainability: environmental, economic, and social.

*True cost accounting* is a method for evaluating the full costs and benefits of systems. For example, applied to different food and farming systems, this means identifying, quantifying, when possible, and making transparent all costs of food production, the prices farmers receive, the affordability of food for consumers, and “externalities” of food production like impacts on the environment and human health. This is also called “full-cost accounting” (Food and Agriculture Organization, 2017). The Economics of Ecosystems and Biodiversity (TEEB) is a comprehensive evaluation framework for doing true cost accounting of food production systems by integrating economic, social, and ecological dimensions and factors through a multidimensional systems approach (Economics of Ecosystems and Biodiversity, 2018a, 2018b). Published on World Environment Day, 5 June 2018, developed and tested over several years at a cost of several million dollars, it is an initiative hosted by the United Nations Environment Program (UN Environment) and coordinated by the TEEB Office in Geneva, Switzerland. It encompasses multiple research and capacity-building projects focusing on the holistic evaluation of agriculture and food systems along their value chains, including their most significant externalities.

Early in 2020, the Trump administration removed Obama-administration regulatory restrictions on polluting rivers and lakes, asserting that such restrictions were too costly to producers. Eco-efficient full-cost accounting would take into account the long-term environmental and human health costs of pollution.

In summary, the *eco-efficient full-cost accounting criterion* calls for a comprehensive, multi-faceted, and holistic approach to assessing costs and benefits. I've described eco-efficiency, triple bottom line, and true cost accounting as examples of evaluation frameworks that illustrate the *eco-efficient full-cost accounting criterion*. True-and-full-cost eco-efficient accounting is, of course, an ideal—and therefore idealistic. But that is one of the purposes of criteria, to define standards and ideals. Addressing this criterion, whether in its complete ideal realization or, more often, in a more modest and limited manner, the point is to think holistically, systemically, and systematically about the costs and benefits of transformational initiatives.

*Eco-efficient full-cost accounting and the coronavirus pandemic.* The pandemic's impact was exacerbated by major shortages of protective gear, ventilators, hospital beds, pharmaceuticals, and, especially in the United States, testing. Mukherjee (2020), an oncologist and author of the best-selling book about the history of cancer, *The Emperor of All Maladies*, has provided an in-depth analysis of **how the short-term efficiency mania in health care administration cost thousands of lives during the pandemic.** He quotes an operations expert at Harvard Business School on the culture of efficiency:

**We've been teaching how to squeeze . . .** Squeeze more efficiency, squeeze cost, squeeze more products at the same cost, squeeze out storage costs, squeeze out inventory. We really need to educate about the value of slack. (p. 30).

Mukherjee asks "To what extent did the market-driven, efficiency-obsessed culture of hospital administration contribute to the crisis?"

His answer: "The numbers in the bean counter's ledger are now body counts in a morgue." (p. 30)

By April, with more than 4 million people sick worldwide and 284,000 Covid-19 deaths, the debate centered on the effects of economic depression versus public health, trying to assess the full costs of the pandemic, a calculation that will go on for some time, a true-cost calculation in which there is vociferous debate about what to include in the full costs. The post-coronavirus economy has become, unintentionally, a transformed economy, "reshaping every aspect of business" (Fortune, 2020), with the nature and extent of the transformation still unfolding.

*Global Alliance for the Future of Food application.* When the Global Alliance was initially formed, the first initiative undertaken was support for developing a true-cost accounting approach to food and agricultural systems. That work became TEEB (Economics of Ecosystems and Biodiversity, 2018a, 2018b). Indeed, to date, the most important initiative of the Global Alliance, one that continues, has been TEEB. **At the heart of the Global Alliance, strategy is employing comprehensive, holistic, ecosystems-based true-cost accounting as a means for catalyzing transformation, an explicit result of transformational engagement, and a characteristic of a transformed food and agriculture system.**

### **Adaptive Sustainability Criterion**

Evaluate transformational sustainability as manifesting ecosystem resilience and adaptability at the nexus between humans and the environment.



### Evaluation implications.

Employ a dynamic view of sustainability.

*Make ecosystem viability and resilience the focus of sustainability not program, project, or intervention continuity.*

### Explanation and elaboration.

Climate change, water shortages, and other environmental crises are bringing home the message loud and clear: We need to connect the dots between human actions across the landscape and seascape, or the earth will cease to care for us. It will cease to grow food, to store water, to host fish and pollinators, to provide energy, medicine and timber. *Changing temperatures will stress systems already overwhelmed by unsustainable patterns of production and consumption, while a growing middle class will further strain planetary boundaries.*

*Many of the solutions however will require breaking down the walls of specific sectors—forestry, agriculture, energy, transport, health—and working with a variety of stakeholders across landscapes, seascapes and cities to achieve multiple goals at once.*

There simply isn't enough time or money to pursue isolated and contradictory solutions. The world is getting smaller—more constrained and interconnected. We have an opportunity to apply system-wide thinking and leverage data to solve the challenges of our time.

Paula Caballero (2015), Senior Director of the World Bank Environment and Natural Resources Global Practice (at the time of this quote)

*Adaptive ecological sustainability has emerged as a priority criterion for evaluation* (Junes, 2019a; Ofir, 2018a, 2018c; Rowe, 2019; Uitto, 2019). Neither the original nor the revised DAC criteria address adaptive ecological sustainability as a priority. The DAC sustainability criterion focuses on continuity: Will intervention outcomes attained last? This criterion is quite understandable from a funder perspective. Funders want to see change and want those changes to be maintained. Evaluators are commissioned to determine both whether the desired and intended changes occurred, and if so, whether they can be continued and sustained. This is fundamentally an accountability perspective imposed from the perspective of funders who must demonstrate that they have made good use of assets entrusted to them. This conceptualization of sustainability as continuation is linear, mechanistic, and static in formulation and evaluation. The logic is as follows: A baseline problem state is identified (needs assessment), an intervention is designed and implemented, the problem is solved (needs met, outcomes achieved), and the solution endures. It is a logic of moving from one condition (a problem) to a new condition (a solution) in a way that the problem does not recur and the solution lasts. This “normal evaluation” paradigm is the core conceptualization of change. It is how evaluators have come to think and practice, and our enforcement and reinforcement of this way of conceptualizing and evaluating sustainable change is a fundamental barrier to realizing an ecological conceptualization of sustainability centered on adaptability and resilience. *Sustainability as adaptive resilience is dynamic, complex, and developmental in both formulation and evaluation* (Patton, 2011, chapter 7).

A note on language: I have called this criterion “*adaptive sustainability*” to distinguish it from the DAC criterion called “sustainability.” Because the DAC criterion is really about continuity not adaptive (resilient ecological) sustainability, I felt a need to include the longer label to be clear about the orientation and substance of this criterion. In a different context simply referring to *sustainability* might suffice, or other adjectives might be used (e.g., resilient sustainability). So let me turn back to elaborating *adaptive sustainability*.

Based on work by ecologist C. S. Holling and applied to social systems by Frances Westley (Westley et al., 2006, 2013), *ecosystem resilience* is defined as “the magnitude of disturbance that can be absorbed

before the system changes its structure by changing the variables and processes that control behavior” (Gunderson & Holling, 2002, p. 28). In formulating and studying ecosystem resilience, Gunderson and Holling articulated *strategic criteria* compatible with both resilience and transformation. They then extended those criteria to human systems and institutions. They found that *resilience* had two quite different meanings in the ecological literature based on two different notions about what it means for a system to be stable. The contrasting and, indeed, competing perspectives on resilience point to the tension created between efficiency, stability, and consistency on the one hand and adaptability on the other.

One definition focuses on efficiency, control, constancy, and predictability—all attributes at the core of desires for fail-safe design and optimal performance. Those desires are appropriate for systems where uncertainty is low, but they can be counterproductive for dynamic, evolving systems where variability and novelty result in high uncertainty. The other definition focuses on persistence, adaptiveness, variability, and unpredictability—all attributes embraced and celebrated by those with an evolutionary or developmental perspective. The latter attributes are at the heart of understanding and designing for sustainability. (Gunderson & Holling, 2002, p. 27)

These different perspectives and definitions led Gunderson and Holling to distinguish two fundamentally different ways of thinking about resilience: *engineering resilience* (continuity sustainability) versus *ecosystem resilience* (adaptable sustainability). Engineering resilience has traditionally focused on “stability near an equilibrium steady state, where resistance to disturbance and speed of return to the equilibrium are used to measure the property.” This is the framework that informs the DAC criterion for sustainability. In contrast, *ecosystem resilience*

emphasizes conditions far from any equilibrium steady state, where instabilities can flip a system into another regime of behavior—i.e. to another stability domain. In this case resilience is measured by the magnitude of disturbance that can be absorbed before the system changes its structure by changing the variables and processes that control behavior. (Gunderson & Holling, 2002, pp. 27–28)

*Sustainability as a Universal Evaluation Criterion.* As noted earlier, the theme of the 2019 Conference of the IDEAs in Prague was *Evaluation for Transformative Change*, supported by a publication with that title (van den Berg et al., 2019). At the conclusion of the conference, participants from around the world adopted a “Declaration on Evaluation for Transformational Change.” The Declaration, adopted October 4, 2019 in Prague, opens as follows:

We, the evaluators, commissioners, parliamentarians and other evaluation users, gathered in the IDEAS Global Assembly and the Third International Conference on Evaluating Environment and Development, recognize the need and urgency of systemic change from local to global levels to address the global crises endangering our future. Having discussed the role of evaluation in promoting learning, systemic and transformational change, we agree on the following statements.

#### *Focus on sustainability.*

In all our evaluations, we commit to evaluating for social, environmental and economic sustainability and transformation, including by assessing contextual factors and systemic changes. We commit to assessing and highlighting, in all evaluations, unintended negative social, economic and environmental effects. (Item 6 of 10 in the Declaration; for the full declaration, see IDEAS, 2019)

Pause and consider the global and universal significance of this declaration about sustainability. All evaluations, emphasis on ALL, are mandated to include attention to sustainability, that is, resilient

ecosystem sustainability. The global climate emergency requires action and engagement by everyone everywhere: *All hands on deck*. This isn't about whether intervention outcomes continue. This is about whether the world is transformed. The time has come for evaluators to embed adaptive sustainability in all frameworks for inquiry to match the call for all nonprofits to incorporate sustainability in their missions (Conway, 2019).

Transformation involves multiple, interdependent dimensions of sustainability. This perspective was well-articulated by Caroline Heider (2017), former Director General Evaluation at the World Bank Group, in her reflections on the DAC criteria.

Taken together these dimensions of sustainability—economic, fiscal, environmental, and social—are complex. It will be hard and costly to try to address them systematically in all evaluations. At the same time, we evaluators cannot afford to turn up with empty hands and concerns about missing data. We need to debate how we would evaluate interventions through these lenses of sustainability, see that the right questions are asked during the design of interventions, and incentivize the collection of relevant data.

George Julnes (2019b) spotlighted and affirmed the importance of evaluators supporting transitions to sustainability as a responsibility for managing processes in the public interest:

... the evaluation community is fortunate to be confronting the challenges of evaluating sustainability, as the complexities calling for a requisite variety of perspectives and responses *now* for sustainability will become evident in other areas of evaluation, and ... *will only become more so*. (p. 147).

## TEXTBOX

The global evaluation website *Better Evaluation* has created a special resource area for tracking evaluation approaches to sustainability called **Footprint Evaluation** ([www.betterevaluation.org/themes/footprint\\_evaluation](http://www.betterevaluation.org/themes/footprint_evaluation)).

*Adaptable sustainability and the coronavirus pandemic.* Scientists use the term “zoonosis” to refer to infectious diseases like COVID-19 (caused by the virus SARS-CoV-2) that spread from animals to humans. The coronavirus pandemic has been associated with a market in Wuhan, China, where the flow of wild animals from forest frontiers to urban communities likely provided the source of infection. The threat from wild animal diseases is exacerbated as human beings place increased pressures on the Earth's diminishing and threatened biodiverse ecosystems. As humanity encroaches further into nature, people have a greater chance of coming into contact with new pathogens carried by animals, and humanity finds itself at greater risk of pandemics. Environmental scientists and infectious disease epidemiologists have determined that “the problems of pandemics, climate and biodiversity loss are linked” (Oakes et al., 2020).

As noted in the opening of this article, the coronavirus pandemic has provided a glimpse into the magnitude of changes set in motion by the looming global climate emergency. New York University climate economist Gernot Wagner has likened the pandemic to “climate change at warp speed” (quoted by Oakes et al., 2020). The evaluation community has become mobilized by the pandemic. That mobilization needs to carry over and magnify by taking seriously and applying the criterion of adaptive sustainability to all evaluations.

*Global Alliance for the Future of Food application.* The first principle adopted by the Global Alliance was **renewability**:

Address the integrity of natural and social resources that are the foundation of a healthy planet and future generations in the face of changing global and local demands.

The Global Alliance was formed on the premise that current systems of industrial food production are unsustainable. The true-cost accounting (TEEB) framework discussed earlier is about a transformation to sustainable food and agricultural systems where sustainability includes human and natural systems together, a resilient ecological perspective defined by the principle of renewability.

### Diversity/Equity/Inclusion (DEI) Criterion

Evaluate how transformational engagement manifests the values of diversity, equity, and inclusion together.

*Evaluation implications.* Evaluate whether, how, and to what extent transformational engagement enhances systems-level diversity, equity, and inclusion.

#### *Explanation and elaboration.*

Wealthy nations and wealthy people use vastly more energy and emit far more carbon dioxide than the poor. . . . The consequences have been wealthy, carbon-intensive lifestyles for some, and a carbon-filled atmosphere for all. (Ellis, 2018, p. 133)

The 21st century has witnessed a growing concentration of wealth and increasing economic inequality. The wealth of 62 people is equal to the wealth of the poorest 3.5 billion people, and the richest 1% have more wealth than the other 99%. Humanity as a whole has had much less impact on Earth than the 1% who acquires more than 80% of the world's wealth generated in a year (Oxfam, 2020).

It is in this context that DEI has become a major thrust of philanthropic engagement, nonprofit advocacy, and evaluation commitment. Indeed, a visible and prominent group of evaluators committed to social justice have long emphasized the principles of diversity, equity, and inclusion. But only recently has DEI become an official value set of philanthropic foundations and nonprofits on a large scale. (For examples of DEI statements, see McKnight Foundation, 2019; Ford Foundation, 2018)).

Writing for the *Nonprofit Quarterly*, Ferris (2019) observed:

Equity, diversity, and inclusion is the topic du jour across disciplines and sectors. Writ large, these themes pertain to increasing the access and power of people and population groups who have been treated unfairly—that is, historically excluded, treated differently, or discriminated against.

DEI illustrates the reciprocal relationship between trends and initiatives in the larger society, including the global social justice uprising that emerged in May, 2020, and what evaluators care about, commit to, and build competency in. The corporate sector has become concerned about DEI, including gathering and reporting data on DEI progress:

Jobs focused on fostering DEI grew by 20 percent in 2018, CEOs are now just as likely to be asked about their DEI efforts as their product roadmap, and leaders who don't create inclusive environments are increasingly being held accountable by their employees, the media, and shareholders. (Emerson, 2019)

The *Economist* (2019) published a feature on DEI tips for chief executives.

Calls for transformation flow from two streams, one values-based and visionary, the other crisis-focused and fear-of-calamity-driven. Transformation as a values-based vision flows from the hopes expressed in the Universal Declaration of Human Rights (adopted in 1948) and subsequently in the Declaration of the Rights of the Child (adopted in 1959). Global DEI norms and values are expressed

and codified in the Declaration on the Rights of Indigenous Peoples and the International Women's Bill of Rights. All people, all of humankind, young and old, have the right to food, water, sanitation, security, shelter, respect, and dignity. As expressed in the ambitious SDGs adopted in 2015, entitled Transforming Our World (UN, 2015), **transformation means No One Left Behind** (Segone & Tateossian, 2017). Thus, in this vision, sustainability and equity, combined, are the foundation for transformation. This links sustainability to equity and transformation. For example, Amnesty International established as its top priority for 2020 tackling the climate crisis by supporting a "human rights-centered transition to a green economy."

Caroline Heider (2017), as former Director General Evaluation at the World Bank Group, has considered this criterion and its implications in depth:

Although the [OECD-DAC] evaluation criteria appear to be neutral and should be applied as such, they were informed by a set of values. The post-2015 agenda has declared its intention to be more inclusive, respecting underprivileged groups of people, which means we as evaluators need to reflect whether the criteria suit these intentions. Being able to shape norms that are more inclusive of diversity rather than judge everyone through more limiting norms will be a necessity if 2030 is to become the world we want. (p. 5)

United Nations Children's Fund and other international agencies have promoted equity-focused evaluation based on human rights and the rights of children (Bamberger & Segone, 2011). This vision for evaluation's role in the world was articulated in the theme of the 2014 Annual Conference of the AEA: Visionary Evaluation for a Sustainable, Equitable Future. Two important evaluation thought leaders, Stewart Donaldson and Robert Picciotto (2016) edited a book on *Evaluation for an Equitable Society*.

**The Equitable Evaluation Initiative (EEI)** promotes the use of evaluation as a tool for advancing equity (TCC Group, 2019). Equitable evaluation encourages evaluators to consider four aspects in their evaluation practice, all at once: **diversity of evaluation teams (beyond ethnic and cultural), cultural appropriateness and validity of evaluation methods, ability of evaluation designs to reveal structural and systems-level drivers of inequity, and the degree to which those affected by what is being evaluated have the power to shape and own how evaluation happens.** (Coffman, 2018; Dean-Coffey, 2018, 2019, 2020, EEI, 2020)

The DEI criterion can include any or all of several such important perspectives:

- Mertens's (1999, 2009) *transformative evaluation paradigm* aimed at ensuring equity for diverse voices of people historically marginalized (Mertens & Wilson 2018).
- Dealing with racism and White privilege, including White frames in evaluation language (Johnson, 2019; Shanker, 2019a, 2019b).
- *Culturally responsive evaluation* (Bowman et al., 2015; Chouinard & Cram, 2019; Hood et al., 2015).
- The *Navigating the Intersection of Culture and Evaluation* framework addresses culture at national or transnational levels to take into account "societal or national dispositions rather than one single culture" (Ofir 2018d, p. 1).
- DEI concerns inclusion of diverse people and perspectives from the Global South in pursuit of global equity, for example, Made in Africa Evaluation (Ofir, 2018b) and incorporating traditional knowledges in climate change initiatives (CTKW, 2014).
- *Decolonizing evaluation*. DEI includes decolonizing both development initiatives and correspondingly decolonizing evaluation (Chouinard & Hopson, 2016; McKegg, 2019). Decolonizing methodologies (Smith, 1999) aim to redress inequities and misrepresentations manifest in research on indigenous peoples and evaluation of programs targeted at indigenous populations.

Evaluation as a profession suffers its own history of racism and White supremacy. Going blue (Blue Marble) and green (environmental) does not exempt us from dealing with Black, Brown, and White. Quite the contrary. Decolonizing evaluation (Chouinard & Hopson, 2016), culturally responsive, and equitable

evaluation have to be part of an evaluation commitment to and engagement with sustainability for human survival on Earth. So, concern for sustainability of the Earth and humanity is connected to DEI.

*Equity and the coronavirus pandemic.* The Covid-19 pandemic has fully exposed the huge health disparities related to poverty and race. The mortality rate from the virus has been twice as deadly for African Americans and Hispanics compared to Whites in the United States (*NY Times*, 2020). The intensity of the social justice uprising following the murder of George Floyd was magnified by disparate economic and social effects of the pandemic.

*Global Alliance for the Future of Food application.* The Global Alliance consists of philanthropic foundations, many of which have adopted DEI within their own foundations. **The Global Alliance has formulated DEI principles that have a transformational sustainability thrust.**

- **Diversity:** Value our rich and diverse agricultural, ecological, and cultural heritage.
- **Equity:** Promote sustainable livelihoods and access to nutritious and just food systems.
- **Inclusion:** Ensure meaningful and authentic engagement of diverse people and organizations in transparent deliberations, shared power, democratic decisions, and collective actions affecting food systems for the public good.

**The Global Alliance principles emphasize the interconnections among these criteria.** That brings us to the final transformation criterion: *interconnectedness momentum*.

### **Interconnectedness Momentum Criterion**

Identify, understand, and evaluate the interconnections that are essential and integral to transformation.

**Evaluation implications.** Evaluate whether, how, and to what extent interconnections among people, networks, institutions, ideas, and movements are deepened and enhanced to support, nurture, catalyze, and accelerate transformational trajectories.

Evaluate whether, how, and to what extent dysfunctional and constraining interconnections are disrupted and broken to liberate positive transformational energy and momentum.

#### *Explanation and elaboration.*

When we try to pick out anything by itself,  
we find it hitched to everything else in the Universe.

John Muir (1838–1914), Pioneering environmentalist

The network is a pattern that is common to all life. Wherever we see life, we see networks. Indeed, at the very heart of the change of paradigm from the mechanistic to the systemic view of life, **we find a fundamental change of metaphor: from seeing the world as a machine to understanding it as a network.**

Capra (2017)

Identifying interconnectedness momentum as a criterion for designing and evaluating transformation is based on the nature of transformation. Systems are defined by interconnections among elements in the system. Transforming systems means changing interconnections within and between systems. **To evaluate interconnections momentum is to evaluate movement toward critical mass and tipping points in transformation theory.** Interconnectedness could be subsumed under the criterion of complex systems framing, but mapping, tracking, understanding, and evaluating interconnections are so

essential to evaluating transformational trajectories that I believe it deserves elevation to a priority criterion for evaluating transformation. This criterion is especially germane to the Global Alliance because its strategy and theory of transformation focus on creating the conditions for and catalyzing interconnections among food systems, agroecological, and climate change networks, researchers, activists, and movements. The evaluation question for this criterion is: To what extent and in what ways are intensifying interconnections generating momentum toward systems transformation?

TIPP (2020) has identified six characteristics of effective social investments to inform financial managers how to address and assess system-level investments for a more sustainable future. One of those characteristics is “contribution to alignment of key stakeholders,” a manifestation of interconnectedness.

### *A Blue Marble Perspective on Interconnections*

Blue Marble refers to the iconic image of the Earth from space without borders or boundaries, a whole Earth perspective. *Blue Marble Evaluation* (Patton, 2020a) focuses on evaluating global systems transformations. This article opened with the premise that we humans are using our planet’s resources, and polluting and warming it, in ways that are unsustainable. Many people, organizations, and networks are working to ensure the future is more sustainable and equitable. Blue Marble evaluators enter the fray by helping design such efforts, providing ongoing feedback for adaptation and enhanced impact, and examining the results of such interventions and initiatives. Incorporating the Blue Marble perspective means looking beyond nation-state boundaries and across sector, issue, and SDG silos to connect the global and local, integrate the human and ecological, and employ evaluative thinking and methods in working with those trying to bring about global systems transformation. *Blue Marble Evaluation* (Patton, 2020a) aspires to make evaluation part of the sustainability solution rather than part of the global climate emergency problem. This means engaging with the understanding the interconnections among global warming; rising seas; loss of ecosystem diversity; species extinction; human food insecurity; growing inequalities; pollution of air, land, and water; spread of virulent diseases; civil unrest; and displacement of millions of people. These trends are interconnected and must be addressed through their interconnectedness.

*Interconnectedness momentum and the coronavirus pandemic.* The pandemic has cast a spotlight on the importance of a coordinated and integrated response to a global emergency based on solid scientific information openly shared. Breaks and gaps in communications, political polarization, inadequate sharing of information and resources, and a relatively powerlessness global governance system have made abundantly clear that nationalism and the territoriality, timidity, and impotence of both national and international agencies cost time and lives. The Trump administration’s attack on the World Health Organization (WHO) and WHO’s limited capacity to mobilize global collaboration are but two of many examples of how a lack of interconnected and shared response exacerbates a global emergency. Science and evidence, including especially evaluation evidence, should be leading the way toward transformation, but science and politics, rather than being positively interconnected, have been at odds in much, though not all, of the coronavirus response. As historian John M. Barry (2020) has observed: “When you mix politics and science, you get politics.”

Two examples. Christi Grimm, the Inspector General at the U.S. Department of Health and Human Services, an evaluation function, surveyed 343 hospitals and documented major shortages of COVID-19 testing kits and personal protective gear. President Trump attacked the findings as “fake” and fired her. Rick Bright, director of the U.S. federal office overseeing COVID-19 vaccine development was fired for refusing, in his words, to let “politics and cronyism drive decisions,” scientific decisions (*The Week*, May 15, pp. 6–7).

Transformation succeeds where multiple systems align. Transformation is undermined and fails where competing systems battle for control and domination rather than collaborate and integrate, at least that is the premise of the *Blue Marble Evaluation* theory of transformation (Patton, 2020a).

*Global Alliance for the Future of Food application.* The Global Alliance for the Future of Food has identified *interconnectedness* as a fundamental principle of transformation: “Understand the implications of the interdependence of food, people, and the planet in a transition to more sustainable food and agricultural systems.” Agricultural and food systems transformation is deeply connected with energy transformation because we are now using one fifth of our fossil fuels to grow and process food. Healthy, sustainably grown food is connected to human health because hunger, food insecurity, and many chronic diseases—heart disease, stroke, and diabetes, as examples—are linked to our diet. More sustainable farming would contribute significantly to alleviating climate change because an organic soil is a carbon-rich soil, which means that it draws CO2 from the atmosphere and locks it up in organic matter. From the perspective of the Global Alliance, food connects all the SDGs (Rockström & Sukhdev, 2016).

## Criteria for Evaluating Transformation

### Summary and Concluding Comments

By way of review, the central significance of interconnectedness for transformation is reinforced by examining the interconnections among the six transformation criteria offered in this article.

**Exhibit 4.** Criteria for Evaluating Transformation.

Criteria	Guidance	Potential Operational Dimensions and Concepts	
1. Transformation fidelity	<p>Assess the extent to which the realities of transformational change initiatives match transformational aspirations and rhetoric.</p> <p><i>Ensure that what is called transformation constitutes transformation</i></p> <p><i>Evaluate whether and how what is called transformational engagement constitutes a trajectory toward transformation</i></p>	<ul style="list-style-type: none"> <li>● Transformation defined contextually</li> <li>● Theory of transformation</li> <li>● Transformation trajectory</li> <li>● Transformational engagement</li> <li>● Transformation elements</li> </ul>	
2. Complex systems framing	<p>Assess systems transformation using systems thinking principles and complexity concepts</p> <p><i>Ensure that transforming systems is the transformational focus</i></p> <p><i>Apply complex systems understandings and frameworks in evaluating transformation</i></p>	<p>Complexity</p> <ul style="list-style-type: none"> <li>● Emergence</li> <li>● Nonlinearities</li> <li>● Dynamics</li> <li>● Adaptation</li> <li>● Co-creation</li> <li>● Path dependence</li> </ul>	<p>Systems</p> <ul style="list-style-type: none"> <li>● Boundaries</li> <li>● Perspectives</li> <li>● Relationships</li> <li>● Dynamics</li> </ul>

(continued)



**Exhibit 4.** (continued)

Criteria	Guidance	Potential Operational Dimensions and Concepts
3. Eco-efficiency full-cost accounting	<p>Document and assess the full costs and benefits of systems transformations, including economic, social, and environmental dimensions</p> <p><i>Compare the full costs and benefits of baseline versus transformed systems</i></p> <p><i>Evaluate whether, how, and to what extent transformational engagement generates net eco-efficient benefits</i></p>	<ul style="list-style-type: none"> <li>● Triple bottom line</li> <li>● Externalities</li> <li>● Ecological, economic, societal costs, and benefits</li> <li>● Direct and indirect costs and benefits</li> <li>● Transparency</li> </ul>
4. Adaptive sustainability	<p>Evaluate transformational sustainability as manifesting ecosystem resilience and adaptability at the nexus between humans and the environment</p> <p><i>Employ a dynamic view of sustainability</i></p> <p><i>Make the ecosystem viability the focus of sustainability not a program, project, or intervention</i></p>	<ul style="list-style-type: none"> <li>● Resilience</li> <li>● Adaptability</li> <li>● Sustainability</li> <li>● Ecosystem vitality</li> </ul>
5. Diversity/equity/inclusion (DEI)	<p>Evaluate how transformational engagement manifests the values of DEI</p> <p><i>Evaluate whether, how, and to what extent transformational engagement enhances systems level diversity, equity, and inclusion</i></p>	<ul style="list-style-type: none"> <li>● Diversity</li> <li>● Inclusion</li> <li>● Equity</li> </ul>
6. Interconnectedness momentum	<p>Identify, understand, and evaluate the interconnections that are essential and integral to transformation</p> <p><i>Evaluate whether, how, and to what extent interconnections among people, networks, institutions, ideas, and movements are deepened and enhanced to support, nurture, catalyze, and accelerate transformational trajectories</i></p> <p><i>Evaluate whether, how, and to what extent dysfunctional and constraining interconnections are disrupted and broken to liberate positive transformational energy and momentum</i></p>	<ul style="list-style-type: none"> <li>● Interrelationships</li> <li>● Interdependencies</li> <li>● Integration</li> <li>● Alignment</li> <li>● Acceleration</li> <li>● Critical mass</li> <li>● Tipping points</li> </ul>

- *Transformation fidelity* involves examining the connection between transformation rhetoric and reality—evaluating the scale, scope, and pace of actual transformational engagement which means reality-testing the vision against what is being done and accomplished.
- *Complex systems framing* connects system thinking and complexity concepts to define the processes, nature, trajectories, and results of transformational engagements.
- *Eco-efficient full cost accounting* makes transparent the interconnections between an intervention's direct costs and benefits in relation to broader environmental and human/societal systems costs and benefits (economic externalities).
- *Adaptive sustainability* invites evaluation of adaptive capacity interconnections between environmental ecosystems and human systems sustainability over time, the interrelationships between people and nature, and, in the triple bottom line economic framing, the interdependence of people, planet, and profits.
- *DEI* focuses attention on the interconnections between who is engaged in and affected by processes of inclusion and diversity toward transformational aspirations of greater equity.
- *Interconnectedness momentum* calls attention to the transformational implications of aligning and integrating across divisions, silos, differing perspectives, historical divisions, and competing interests toward a vision of a more sustainable and equitable future.

Exhibit 4 presents a summary of these six criteria for evaluating transformation. The middle column presents the definition of each criterion. Each criterion is a cluster concept made up of multiple dimensions. The right-hand column identifies some of the dimensions and factors that may inform the design and evaluation of transformations. In particular contexts for specific evaluations, any of the multiple dimensions of the criteria highlighted here might be elevated to the status of a priority criterion.

### *The Importance of Developing Contextually Appropriate Criteria*

I want to reiterate that the criteria I have offered for evaluating transformation are meant to be illustrative of what is possible. The applications of the criteria to the coronavirus pandemic and the Global Alliance for the Future of Food provide illustrations of contextual meaningfulness rather than suggesting these specific criteria be standardized and adopted as universal. Schwandt (2018) has articulated definitively the case against treating the DAC criteria, or any set of criteria, as universal.

ANY effort to stipulate, authorize, legitimate (whether directly or indirectly) a particular set of criteria as stable or permanent or semi-permanent, etc. is a mistake . . . because criteria (and the values on which they rest) are something to be debated and negotiated in every evaluation. We cannot transcend the limitations, uncertainty, and contingency of our knowing and valuing by positing some set of criteria as THE way we are to evaluate.

Williams (2018) adds that “one of the negative consequences of the DAC criteria is that they have, in many cases, removed the responsibility and . . . ability of many evaluators in the ‘development’ space to be able to develop, understand and work within ‘criteria’.”

More diverse perspectives on criteria can be found in the dialogue stimulated by Ofir's blog posts on the DAC criteria and the energetic and engaged comments and reactions that followed her own insightful blog posts (Ofir, 2018a, 2018c, 2018e). Directly addressing what criteria enable “valuing for development,” Ofir (2018c) distinguishes three different sets of criteria negotiable with stakeholders to varying degrees:

- **Nonnegotiable criteria** determined by the characteristics of development, for example, impact, equity, and sustainability;
- **criteria imposed by institutional mandates**, values and norms, for example, gender-equity and support for human rights; and
- **negotiable criteria** determined by specific stakeholder interests in particular contexts, for example, value for money and scalability.

There is a tension and trade-off between specifying standardized universal evaluation criteria like equity and adaptive sustainability versus supporting entirely situation-specific criteria. I am sympathetic to the desirability of considering and offering alternative criteria situationally and contextually rather than imposing predetermined standardized and purportedly universal criteria. But, I believe it is appropriate and useful to have a set of standardized universal criteria endorsed by professional evaluators (through their associations) as a starting place for considering either adoption or customization. Menus of criteria might be generated to stimulate consideration of what resonates in a particular context. The DAC consultation process generated many suggestions for additional and alternative evaluation criteria. These could be pulled together and organized as a menu of criteria for consideration and adaptation in the context of a particular evaluation. Let us consider what such menu options might include and how they could be used.

### *Menu of Criteria Options in Substance and Labeling*

Feinstein (2019) suggests that “**significance**” be a transformational criterion. This is a different way of expressing the criterion I’ve called transformational fidelity. I ask, “Is the initiative genuinely transformative?” Feinstein asks: Is it transformatively significant?

In practical terms, evaluations that aim to be transformative would need to include not only the standard evaluation criteria (relevance, effectiveness/efficacy, efficiency, sustainability and impact), as well as coherence [but] also significance . . . An intervention (say a policy) may be highly relevant and get high marks on the other [DAC] four criteria, and yet it may not change in any significant way the situation of the target population. Its actual effects, although positive, may be insignificant in terms of transformational change. As this cannot be captured through the “big five” criteria, it may be worthwhile to introduce another criterion, significance, that points to the size of the effect. It should be noted that significance goes well beyond “statistical significance” since it also embraces the size of direct and indirect effects. It is concerned with “practical importance.” (pp. 19–20)

Picciotto (2020a) also advocates significance as a criterion. Drawing on the definition of evaluation as determining merit, worth, and significance, he would make significance the universal criterion for judging the transformative impact of interventions:

[S]ignificance is the end-game when all pertinent data and evaluative judgments, including the size, importance and transformative effects of the social action, are synthesized to reach an overall judgment of value . . . Significance assessment identifies the nature and weighs the extent of the gaps between merit and worth and strikes a judicious balance that seeks to value the social importance and transformative impact of the evaluand. Significance is therefore about doing good as well as doing right from a public interest perspective, with special consideration for the interests of the most disadvantaged groups in society. (p. 55)

When I advocate that criteria be discussed and determined by primary intended users of the evaluation, I’m talking about deepening their understanding and buy-in by determining whether to call the criterion “transformational fidelity,” “significance,” “practical importance,” or something else. The criterion might be called “seriously transformative,” “genuinely transformative,” “authentically transformative,” “majorly transformative,” or whatever language resonates with intended users in that

context and situation. The red thread running through all those labels is that transformation involves major systems change and the evaluation should examine the extent to which major systems change is targeted and accomplished, or at least on a trajectory to be accomplished.

I struggled with what term to use to distinguish transformation-focused sustainability from the DAC definition of sustainability as continuity. I have offered *adaptive sustainability* as a criterion for evaluating transformation. Rowe has suggested “sustainability-ready” as a criterion for judging intervention proposals. I was attracted to the term “sustainment” as “moving beyond sustainability” (Suarez, 2020) because the latter sounds like possibility while the former implies having arrived. Sustainability suggests capacity to achieve sustainment (the future being sustained), but the term is too unfamiliar to be user-friendly.

Likewise, the criterion I have called “interconnectedness momentum” might be called “alignment,” “integration,” “critical mass,” “tipping point trajectory,” or, again, whatever term or phrase resonates within a particular context.

Another example of differential wording preference concerns simply using the term “transformation” versus “transformational change.” For me, “transformational change” is redundant, but for some that phrase resonates more than just “transformation.” For others, the preferences are reversed.

Yet another example: When discussing the complexity criterion, I noted that whether to treat systems thinking and complexity framing as separate criteria, or combine them as I have done in this instance, is part of the decision to be made for any given design and evaluation of transformation depending on what will be most useful in conceptualizing the nature of the transformational engagement and aligned evaluation.

The point of sharing these examples of alternative criteria is to illustrate the potential value of a menu of transformative criteria with explanations and justifications of each, including alternative wording, but not insisting on one set of universally standardized and mandated criteria. In reviewing the menu of criteria options, stakeholders, intended users, and evaluators would be *forced to engage in actual evaluative thinking about criteria*—what they mean, why they exist, how to use them—instead of simply complying with a preordained set of criteria mandated by some authoritative body, such banal compliance by evaluators being a major criticism of how the DAC criteria had come to be used. Perhaps that should be a criterion for judging both interventions and evaluations: **thoughtfulness.**

### **By What Authority?**

One peer reviewer of an earlier draft of this article raised a “concern . . . that it offers no justification for the authority/legitimacy of the proposed criteria. Whose criteria are these? Whose interests do they serve? (Imagine comparing the source of legitimacy and authority of the OECD/DAC criteria to the source of legitimacy/authority of the proposed new criteria.)”

Interestingly, that makes institutional authority a criterion for determining legitimacy and credibility. By that criterion, what is offered here has no credibility. The endnote describes how these proposed criteria for judging transformation were developed and tested. But to the reviewer’s point, they have not been approved by any official body. Indeed, as the foregoing discussion makes clear, in my view, the authority, legitimacy, and credibility of evaluation criteria for judging the quality of an intervention should reside with the stakeholders, intended users, and evaluator(s) engaged in designing, conducting, and using any particular evaluation. These proposed criteria are offered as a deliberative stimulus and menu for consideration and decision making by those who are responsible for the credibility and legitimacy of a given evaluation.

### **Conclusion**

Let me close by offering three conclusions about criteria generally:

1. Evaluation criteria express values and, in so doing, establish “what is good.” As Stame (2018) has posited, “Taking into consideration public values beyond those stated in the program objectives requires identifying appropriate criteria and standards for identifying what ‘ought to be’” (p. 446). Thus, in addition to measuring attainment of program goals and answering specific evaluation questions, evaluators should address what is in the public interest, as mandated by the AEA Guiding Principles. Articulating criteria means taking “an ethical stance, which covers principles of what is good as well as moral expertise” (Stame, 2018, p. 448). To reiterate, criteria have both methodological and ethical implications. Criteria are both technical and moral, both instrumental and values-based, and both descriptive (what to usefully examine) and prescriptive (what ought to matter).
2. Standardized universal evaluation criteria can offer consistent guidance but cannot cover the rich variety of kinds of interventions and evaluations in this marvelously diverse world, so there is substantial value in having the primary intended users of any evaluation engage in a process of reviewing universal criteria options and deliberating what criteria are appropriate for a given evaluation, including what to name the criteria chosen. Departures from or omissions of any universally affirmed evaluation criteria (affirmed by professional evaluation associations like AEA’s Guiding Principles) should be explicitly addressed and justified.
3. Articulating criteria is a necessary but insufficient action for ensuring quality interventions and their appropriate evaluation. Any criteria identified must be applied diligently, implemented systematically, used appropriately, and evaluated for adherence, meaningfulness, and results attained (Patton, 2018).

### *Criteria in the Context of the Global Climate Emergency*

We stand at a critical juncture in our collective efforts to limit dangerous global heating. By the end of the coming decade we will be on one of two paths. One is the path of surrender, where we have sleepwalked past the point of no return, jeopardizing the health and safety of everyone on this planet. Do we really want to be remembered as the generation that buried its head in the sand, that fiddled while the planet burned?

The other option is the path of hope. A path of resolve, of sustainable solutions.

UN Secretary-General António Guterres (2019)

Let me conclude with some reflections about the importance of having explicit criteria for transformation. Evaluators need to engage and take seriously the global emergency—and do so with a sense of urgency and transformational scale. The DAC criteria are intentionally technical and neutral. Relevance, effectiveness, efficiency, impact, continuity (sustainability), and coherence communicate evaluation priorities. They are framed as applicable to any and all interventions. That is both their strength and weakness, their strength in that they posit universal applicability, and their weaknesses in that they are correspondingly milquetoast and bland. They are old news. They fail to inspire. In contrast, if evaluators are to become part of the solution to the global crisis, we must engage with the scale, scope, and urgency of transformation. **The criteria offered here are values-based. They make it clear that evaluators have a stake in transformation and manifest our stake in the future of humanity by addressing transformation explicitly and directly.** The values-based nature of the transformation criteria will be a source of controversy, even disdain, for those who prefer to treat evaluation as an independent technical activity. Transformation-focused criteria treat evaluation as part of the transformation process and evaluators as having skin in the game (Patton, 2020), the game being the future of humanity on Earth.

Distinguished management consultant Peter Drucker asserted that:

The greatest danger in times of turbulence is not the turbulence—it is to act with yesterday's logic. (quoted by Carrigan, 2010, p. 99)

Applied to evaluation this becomes:

The greatest danger for evaluators in times of turbulence is not the turbulence—it is to act with yesterday's logic and criteria.


### 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.

### ORCID iD

Michael Quinn Patton  <https://orcid.org/0000-0002-4706-2941>

### Note

1. *Consultative process for transformation criteria*: In the summer of 2019, I had the opportunity to teach a course on Global Systems Transformation Evaluation (*Blue Marble Evaluation*, explained below) at the IPDET in Bern, Switzerland. We devoted considerable time to reviewing the Development Assistance Committee (DAC) criteria and considering alternative criteria for evaluating transformation. I shared those results at the International Development Evaluation Association meeting in Prague in October, 2019, where I got further feedback. The discussion stimulated by Ofir's blog posts on the DAC criteria provided another source for considering both the DAC criteria and alternatives for transformation. Deborah Rugg facilitated leadership training for transformation in the United Nations system and chaired a session on the topic at the 2018 American Evaluation Association Annual Conference, both of which provided opportunities to consider principles and criteria for evaluating transformation. In writing the *Blue Marble Evaluation* book about evaluating global systems transformation (Patton, 2020), the Blue Marble Evaluation team engaged in developing principles and criteria for evaluating transformation. Thus, the criteria offered here are the results of 2 years of reflection, consultation, workshopping possibilities, and feedback about evaluating criteria for transformation. In sharing them here, let me emphasize that I mean for them to illustrate possibilities and stimulate further contextual adaptation not to be treated as universal, standardized, and/or mandated criteria.

### References

- American Evaluation Association. (2018). Guiding principles for evaluators. <https://www.eval.org/p/cm/ld/fid=51>
- Azzam, T. (2018). The application of the logic of evaluation to the real world: Understanding the four-step logic of evaluation process. *Psychology Today*. <https://www.psychologytoday.com/us/blog/evaluation-social-betterment/201810/the-application-the-logic-evaluation-the-real-world>
- Bamberger, M., & Mabry, L. (2020). *Realworld evaluation* (3rd ed.). Sage.
- Bamberger, M., & Segone, M. (2011). *How to design and manage equity-focused evaluations*. UNICEF.
- Bamberger, M., Vaessen, J., & Raimondo, E. (2016). *Dealing with complexity in development evaluation: A practical approach*. Sage.
- Barry, J. M. (2020, May 1). Politics and science. *The Week*, p. 17.

- Better Evaluation. (2020). *Adapting evaluation in the time of COVID-19*. <https://www.betterevaluation.org/en/blog/adapting-evaluation-time-covid-19-part-1-manage>
- Bitar, K. (2020, May 1). If not now, when? Evaluators as agents of change in the time of a global pandemic. *May 1 Blog*. <http://khalil-bitar.com/index.php/blog/>
- Bowman, N. R., Francis, C. D., & Tyndall, M. (2015). Culturally responsive indigenous evaluation a practical approach for evaluating indigenous projects in tribal reservation contexts. In S. Hood, R. Hopson, & H. Frierson (Eds.), *Continuing the journey to reposition culture and cultural context in evaluation theory and practice* (pp. 335–359). Information Age.
- Capra, F. (2004). *The hidden connections: Integrating the biological, cognitive, and social dimensions of life into a science of sustainability*. Anchor.
- Capra, F. (2017, February 5). The science of connection. *Uplift Blog*. <https://upliftconnect.com/science-of-connection/>
- Cartier, K. M. S. (2020, April 24). Eight lessons from COVID-19 to guide our climate response. EOS (science news) blog. <https://eos.org/articles/eight-lessons-from-covid-19-to-guide-our-climate-response>
- CDC (Centers for Disease Control and Prevention) (2020) Flattening the coronavirus curve. *New York Times*, March 27. <https://www.nytimes.com/article/flatten-curve-coronavirus.html>
- Chaplowe, S. (2020, May 4). Evaluation in the face of the climate emergency: Observations and synthesis. *Blue Marble Evaluation Blog*.
- Chelsky, J., & Kelly, L. (2020, April 1). Bowling in the dark: Monitoring and evaluation during COVID-19 (coronavirus): Lessons from past experience can help creatively and responsibly adapt M&E practices. *Independent Evaluation Group Blog*. <https://ieg.worldbankgroup.org/blog/mande-covid19>
- Chouinard, J. A., & Cram, F. (2019). *Culturally responsive approaches to evaluation: Empirical implications for theory and practice*. Sage.
- Chouinard, J., & Hopson, R. (2016). Decolonizing international development evaluation. *Canadian Journal of Program Evaluation*, 30(3, December).
- Christensen, C. M., Raynor, M. E., & McDonald, R. (2015, September). What is disruptive innovation? Harvard Business Review.
- Climate and Traditional Knowledges Workgroup (CTKW). (2014). Guidelines for considering traditional knowledges in climate change initiatives. <http://climatetkw.wordpress.com/>
- Coffman, J. (2018). *Equitable evaluation applies to all evaluation*. Center for Evaluation Innovation. <https://www.evaluationinnovation.org/insight/equitable-evaluation-applies-to-all-evaluation/>
- Conway, M. (2019, February 11). Put climate in every nonprofit mission. *Nonprofit Quarterly Blog*.
- DAC Network on Development Evaluation. (2018). OECD DAC evaluation criteria: Summary of consultation responses. [https://ieg.worldbankgroup.org/sites/default/files/Data/DAC-Criteria/ConsultationReport\\_EvaluationCriteria.pdf](https://ieg.worldbankgroup.org/sites/default/files/Data/DAC-Criteria/ConsultationReport_EvaluationCriteria.pdf)
- Davidson, E. J. (2005). *Evaluation methodology basics: the nuts and bolts of sound evaluation*. Sage.
- Dean-Coffey, J. (2018). What's race got to do with it? Equity and philanthropic evaluation practice. *American Journal of Evaluation*, 39(4), 527–542.
- Dean-Coffey, J. (2019). Evaluation in an equitable society: one person's perspective. *Evaluation and Program Planning*, 75, 89–90.
- Dean-Coffey, J. (2020). How a seasoned consultant navigates social-justice-oriented evaluation. *AEA365 blog*, January 22. [https://aea365.org/blog/how-a-seasoned-consultant-navigates-social-justice-oriented-evaluation-by-jara-dean-coffey/?utm\\_source=feedburner&utm\\_medium=email&utm\\_campaign=Feed%3A+aea365+%28AEA365%29](https://aea365.org/blog/how-a-seasoned-consultant-navigates-social-justice-oriented-evaluation-by-jara-dean-coffey/?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+aea365+%28AEA365%29)
- Donaldson, S. I., & Picciotto, R. (Eds.). (2016). *Evaluation for an equitable society*. Information Age.
- Duhigg, C. (2020, May 4). Seattle's leaders let scientists take the lead. New York's did not. *The New Yorker*, pp. 16–22.
- The Economics of Ecosystems and Biodiversity. (2018a). *Measuring what matters in agriculture and food systems: A synthesis of the results and recommendations of TEEB for Agriculture and Food's Scientific and Economic Foundations report*. UN Environment.

- The Economics of Ecosystems and Biodiversity. (2018b). *TEEB for agriculture & food: Scientific and economic foundations*. UN Environment. <http://teebweb.org/agrifood/home/scientific-and-economic-foundations-report/>
- EI (Equitable Evaluation Initiative) (2020). *Equitable evaluation initiative*. <https://www.equitableeval.org/>
- Ehrenfeld, J. R. (2005). "Eco-efficiency: Philosophy, theory and tools." *Journal of Industrial Ecology*, 9(4), 6–8. <https://doi.org/10.1162/108819805775248070>
- Elkington, J. (2018, June 25). 25 years ago i coined the phrase "triple bottom line." Here's why it's time to rethink it. *Harvard Business Review*.
- Ellis, E. C. (2018). *Anthropocene*. Oxford University Press.
- Emerson, J. (2019). What's in store for corporate diversity, equity, and inclusion in 2019? *Paradigm blog*, January 30. <https://medium.com/inclusion-insights/what-in-store-for-corporate-diversity-equity-and-inclusion-in-2019-d3c1292dd920>
- Euber, J. (2020, May 1). Unanticipated crisis lessons: Nonprofit's pivot to online increases participation. *NPQ Blog*. [https://nonprofitquarterly.org/unanticipated-crisis-lessons-nonprofits-pivot-to-online-increases-participation/?utm\\_source=NPQ+Newsletters&utm\\_campaign=95879bfe53-EMAIL\\_CAMPAIGN\\_2018\\_01\\_11\\_COPY\\_01&utm\\_medium=email&utm\\_term=0\\_94063a1d17-95879bfe53-12318013&mc\\_cid=95879bfe53&mc\\_eid=707ec28f50](https://nonprofitquarterly.org/unanticipated-crisis-lessons-nonprofits-pivot-to-online-increases-participation/?utm_source=NPQ+Newsletters&utm_campaign=95879bfe53-EMAIL_CAMPAIGN_2018_01_11_COPY_01&utm_medium=email&utm_term=0_94063a1d17-95879bfe53-12318013&mc_cid=95879bfe53&mc_eid=707ec28f50)
- Eval4action (2020). The decade of evaluation for action. <https://www.eval4action.org/>
- Fazey, I. (2020). Keynotes on transformation and sustainability. <https://ioanfazey.com/keynotes/>
- Fazey, I., Schöpke, N., Caniglia, G., Patterson, J., Hultman, J., van Mierlo, B., Säwe, F., Wiek, A., Wittmayer, J., Aldunce, P., Waer, H. A., Battacharya, N., Bradbury, H., Carmen, E., Colvin, J., Cvitanovic, C., D'Souza, M., Gopel, M., Goldstein, B., & Hämäläinen, T. (2018). Ten essentials for action-oriented and second order energy transitions, transformations and climate change research. *Energy Research and Social Science*, 40, 54–70.
- Feinstein, O. (2019). Dynamic evaluation. In R. B. van den Berg, C. Magro, & S. S. Mulder (Eds.), *Evaluation for transformational change: Opportunities and challenges for the sustainable development goals* (pp. 17–31). IDEAS. [https://ideas-global.org/wp-content/uploads/2019/11/2019-11-05-Final\\_IDEAS\\_EvaluationForTransformationalChange.pdf](https://ideas-global.org/wp-content/uploads/2019/11/2019-11-05-Final_IDEAS_EvaluationForTransformationalChange.pdf)
- Feinstein, O. (2020, April 28). Dynamic evaluation. *Evaluation for Development Blog*. <https://zendaofir.com/dynamic-evaluation-for-societal-transformation/>
- Food and Agriculture Organization. (2017). *Sustainability pathways*. <http://www.fao.org/nr/sustainability/full-cost-accounting/e>
- Ford Foundation (2018) Diversity, equity, inclusion. <https://www.fordfoundation.org/about/people/diversity-equity-and-inclusion/>
- Forss, K., Marra, M., & Schwartz, R. (Eds.). (2011). *Evaluating the complex: Attribution, contribution, and beyond* (Comparative Policy Evaluation Series, vol. 18). Transaction.
- Fortune. (2020, May). The coronavirus economy, how the pandemic is reshaping every aspect of business. *Fortune Magazine*.
- Furubo, J.-E., Rist, R. C., & Speer, S. (Eds.). (2013). *Evaluation and turbulent times: Reflections on a discipline in disarray*. Transaction.
- Global Alliance for the Future of Food (2020). The Global Alliance formally adopts a theory of transformation. <https://futureoffood.org/the-global-alliance-makes-history-with-formal-adoption-of-a-theory-of-transformation/>
- Global Environment Facility (2018). *Strengthening monitoring and evaluation of climate change adaptation*. <https://www.thegef.org/publications/strengthening-monitoring-and-evaluation-climate-change-adaptation>
- Gunderson, L. H., & Holling, C. S. (2002). *Panarchy: understanding transformations in human and natural systems*. Island Press.
- Guterres, A. (2019, December). UN Secretary-General remarks at opening ceremony of UN Climate Change Conference COP25.



- Guterres, A. (2020, April 28). Collaborative leadership is critical to fight both coronavirus and climate change. *TIME*. <https://time.com/5828535/un-chief-coronavirus-climate-change-collaboration/>
- Hayes, D. J. (2020). *Flattening the global warming curve*. NYU Energy & Environmental Impact Center.
- Heider, C. (2017). *Rethinking evaluation*. The World Bank Independent Evaluation Group. <http://ieg.worldbankgroup.org/sites/default/files/Data/RethinkingEvaluation.pdf>
- Hepburn, C., O'Callaghan, B., Stern, N., Stiglitz, J., & Zenghelis, D. (2020). Will COVID-19 fiscal recovery packages accelerate or retard progress on climate change? Smith School Working Paper 20-02. <https://www.smithschool.ox.ac.uk/publications/wpapers/workingpaper20-02.pdf>
- Hernandez-Licona, G. (2019). Foreword. In R. B. van den Berg, C. Magro, & S. S. Mulder (Eds.), *Evaluation for transformational change: Opportunities and challenges for the sustainable development goals* (pp. x–xii). IDEAS, Exeter. [https://ideas-global.org/wp-content/uploads/2019/11/2019-11-05-Final\\_IDEAS\\_EvaluationForTransformationalChange.pdf](https://ideas-global.org/wp-content/uploads/2019/11/2019-11-05-Final_IDEAS_EvaluationForTransformationalChange.pdf)
- Hodgson, A. (2019, November 28). Transformation needs fresh perspectives [Blog post]. <https://www.transformationsforum.net/transformation-needs-fresh-perspectives/>
- Hodgson, A. (2020). *Systems thinking for a turbulent world: A search for new perspectives*. Routledge.
- Hood, S., Hopson, R., & Frierson, H. (Eds.). (2015). *Continuing the journey to reposition culture and cultural context in evaluation theory and practice*. Information Age.
- Hussain, S., & Vause, J. (2018). TEEB for agriculture & food: Background and objectives. In S. Hussain & J. Vause (Eds.), *TEEB for agriculture & food: Scientific and economic foundations* (pp. 3–5). UN/FAO.
- Independent Evaluation Group. (2016). *Supporting transformational change for poverty reduction and shared prosperity—Lessons from the World Bank experience*. World Bank.
- Independent Evaluation Group. (2018). *Evaluation of GEF support for transformational change*. Global Environmental Facility. <file:///C:/Users/Owner/Documents/Transformation/DAC%20criteria/transformational-engagement-2017-brief.pdf>
- Independent Evaluation Group. (2020a). *Evaluative resources and lessons to inform the COVID-19 response*. The World Bank: IEG. [http://ieg.worldbankgroup.org/topic/covid-19-coronavirus-response?utm\\_source=Independent+Evaluation+Group+Contacts&utm\\_campaign=fee9c85338-EMAIL\\_CAMPAIGN\\_2020\\_05\\_04\\_05\\_44\\_COPY\\_03&utm\\_medium=email&utm\\_term=0\\_fc6e7f2a32-fee9c85338-113584913](http://ieg.worldbankgroup.org/topic/covid-19-coronavirus-response?utm_source=Independent+Evaluation+Group+Contacts&utm_campaign=fee9c85338-EMAIL_CAMPAIGN_2020_05_04_05_44_COPY_03&utm_medium=email&utm_term=0_fc6e7f2a32-fee9c85338-113584913)
- Independent Evaluation Group. (2020b, April 1). Lessons from past experience can help creatively and responsibly adapt M&E practices. *Independent Evaluation Group Blog*. <https://ieg.worldbankgroup.org/blog/mande-covid19>
- The Investment Integration Project. (2020, May 12). Who is doing well, who isn't: Evaluating the performance of asset managers against systemic social and environmental progress. Online Panel Discussion.
- International Development Evaluation Association. (2019). *Prague declaration on evaluation for transformational change*. International Development Evaluation Association, adopted on Friday, October 4. <https://ideas-global.org/wp-content/uploads/2019/10/Prague-Declaration-4-October-2019.pdf>
- Johnson, A. F. (2019). (On the absence of) whiteness. *AEA365 blog*, June 25. [https://aea365.org/blog/a-look-at-language-week-on-the-absence-of-whiteness-by-a-rafael-johnson/?utm\\_source=feedburner&utm\\_medium=email&utm\\_campaign=Feed%3A+aea365+%28AEA365%29](https://aea365.org/blog/a-look-at-language-week-on-the-absence-of-whiteness-by-a-rafael-johnson/?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+aea365+%28AEA365%29)
- Joint Committee on Standards for Educational Evaluation. (2010). *The program evaluation standards*. <http://www.jcsee.org/program-evaluation-standards>
- Julnes, G. (Ed.), (2019a). Evaluating sustainability: Evaluative support for managing processes in the public interest. *New Directions for Evaluation*, 162 (Summer).
- Julnes, G. (2019b). Supporting transitions to sustainability: Evaluation for managing processes in the public interest. In G. Julnes (Ed.), *New Directions for Evaluation*, 162, 119–154.
- Karalisi, T. (2020). Planning and evaluation during educational disruption: Lessons learned from covid-19 pandemic for treatment of emergencies in education. *European Journal of Education Studies*, 7(4), <https://oapub.org/edu/index.php/ejes/article/viewFile/3047/5685>

- Kennedy-Chouane, M. (2020). *Evaluation criteria: What has changed, why, and why does it matter?* UNICEF/EVALSDGS webinar, January 16. <https://www.youtube.com/watch?v=r3jzS16D-QM&feature=youtu.be>
- Kleissner, C. (2020, January 16). *Impact investing for systems change*. Impact Entrepreneur's Luminarias Series.
- Kolbert, E. (2020, January 13). Don't wait. *The New Yorker*, pp. 13–14.
- Lundgren, H. (2019). *OECD evaluation norms and criteria*. <https://www.mofa.go.jp/mofaj/gaiko/oda/files/000447961.pdf>
- Lydenberg, S., & Burckart, B. (2020). *Assessing system-level investments: A guide for asset owners*. The Investment Integration Project. [https://www.tiiproject.com/wp-content/uploads/2020/04/Assessing-System-Level-Investments\\_FINAL\\_04-21-2020.pdf](https://www.tiiproject.com/wp-content/uploads/2020/04/Assessing-System-Level-Investments_FINAL_04-21-2020.pdf)
- McKegg, K. (2019). White privilege and the decolonization work needed in evaluation to support indigenous sovereignty and self-determination. *Canadian Journal of Program Evaluation*, 34(2, Fall), 357–367. <https://doi.org/10.3138/cjpe.67978>
- McKnight Foundation (2019). *Diversity, Equity, Inclusion*. <https://www.mcknight.org/news-ideas/mcknight-foundation-statement-diversity-equity-inclusion/>
- Mertens, D. M. (1999). Inclusive evaluation: Implications of transformative theory for evaluation. *American Journal of Evaluation*, 20(1), 1–14.
- Mertens, D. M. (2009). *Transformative research and evaluation*. Guilford
- Mertens, D. M., & Wilson, A. T. (2018). *Program evaluation theory and practice* (2nd ed.). Guilford.
- Miller, R. L., & Campbell, R. (2006). Taking stock of empowerment evaluation: An empirical review. *American Journal of Evaluation*, 27(3), 296–319.
- Moore, G. F., Evans, R. E., Hawkins, J., Littlecott, H., Melendez-Torres, G. J., Bonell, C., & Murphy, S. (2019). From complex social interventions to interventions in complex social systems: Future directions and unresolved questions for intervention development and evaluation. *Evaluation*, 25(1), 23–45. <https://doi.org/10.1177/1356389018803219>
- Mora, C., Spirandelli, D., Franklin, E. C., Lynham, J., Kantar, M. B., Miles, W., Smith, C. Z., Freel, K., Moy, J., Louis, L. V., Barba, E. W., Bettinger, K., Frazier, A. G., Colburn, I. X. J. F., Hanasaki, N., Hawkins, E., Hirabayashi, Y., Knorr, W., Little, C. M., & Emanuel, K. (2018). Broad threat to humanity from cumulative climate hazards intensified by greenhouse gas emissions. *Nature Climate Change*, 8, 1062–1071.
- Mukherjee, S. (2020, May 4). After the storm: The pandemic has revealed dire flaws in American medicine. *The New Yorker*, 24–31.
- Müller, A., & Sukhdev, P. (2018). Preface. In *The economics of ecosystems and biodiversity* (TEEB). (2018). *Measuring what matters in agriculture and food systems: A synthesis of the results and recommendations of TEEB for Agriculture and Food's Scientific and Economic Foundations Report* (pp. 3–5). UN Environment.
- Natsios, A. (2010). *The clash of the counter-bureaucracy and development*. Center for Global Development. <https://www.cgdev.org/publication/clash-counter-bureaucracy-and-development>
- NY Times*. (2020, April 14). Black Americans face alarming rates of coronavirus infection. <https://www.nytimes.com/2020/04/07/us/coronavirus-race.html>
- Oakes, L. E., Olson, S. H., & Watson, J. (2020, April 22). Preventing pandemics, global warming and environmental degradation all at once. *CNN*. <https://www.cnn.com/2020/04/22/opinions/pandemics-global-warming-linked/index.html>
- Ofir, Z. (2018a, February 15). Updating the DAC evaluation criteria, part 5. Non-negotiable criteria. *Evaluation for Development Blog*. <http://zendaofircom/updates-dac-evaluation-criteria-part-5/>
- Ofir, Z. (2018b, June 8). Made in Africa evaluation, part 2: “Africa-rooted” evaluation. *Evaluation for Development Blog*.
- Ofir, Z. (2018c, July 31). *Evaluation for Development Blog*. <http://zendaofircom/zendas-top-ten-tips-for-yees-2/>
- Ofir, Z. (2018d, August 28). The NICE framework, part 1. Where culture and evaluation meet. *Evaluation for Development Blog*.

- Ofir, Z. (2018e, October 13). Updating the DAC criteria. *Evaluation for Development Blog*. <http://zendaofir.com/dac-criteria-part-11/>
- Ofir, Z. (2020, March 29–April 20). Transforming evaluations and COVID-19. *Evaluation for Development* [4 blog posts]. <https://zendaofir.com/transforming-evaluations-and-covid-19-part-4-accelerating-change-in-practice/>
- Organization for Economic Cooperation and Development. (2013). The DAC network on development evaluation: 30 years of strengthening learning in development. OECD DAC Network on Development Evaluation.
- Organization for Economic Cooperation and Development. (2020). Evaluation criteria frequently asked questions. <https://www.oecd.org/dac/evaluation/eval-criteria-faq.htm>
- Organization for Economic Cooperation and Development Secretariat. (2002). Indicators to measure decoupling of environmental pressure from economic growth. Sustainable development. [http://www.docstoc.com/docs/84838188/oecd\\_decoupling](http://www.docstoc.com/docs/84838188/oecd_decoupling)
- Organization for Economic Cooperation and Development/Development Assistance Committee Network on Development Evaluation. (2019a). *Better criteria for better evaluation: Revised evaluation criteria definitions and principles for use*. Author. [www.oecd.org/dac/evaluation](http://www.oecd.org/dac/evaluation)
- Organization for Economic Cooperation and Development/Development Assistance Committee Network on Development Evaluation (2019b). *Evaluation criteria*. <https://www.oecd.org/dac/evaluation/daccriteriafor-evaluatingdevelopmentassistance.htm>
- Oxfam (2020). *World's billionaires have more wealth than 4.6 billion people*. <https://www.oxfam.org/en/press-releases/worlds-billionaires-have-more-wealth-46-billion-people>
- Ozeki, S., Coryn, C. L. S., & Schröter, D. C. (2019, October). Evaluation logic in practice: Findings from two empirical investigations of American Evaluation Association members. *Evaluation and Program Planning*, 76. <https://www.sciencedirect.com/science/article/abs/pii/S0149718919301429>
- Parsons, B., Dhillon, L., & Keene, M. (2020). *Visionary evaluation for a sustainable, equitable future*. Information Age.
- Patton, M. Q. (2008). *Utilization-focused evaluation*. Sage.
- Patton, M. Q. (2011). *Developmental evaluation: applying complexity concepts to enhance innovation and use*. Guilford.
- Patton, M. Q. (2015). *Essentials of utilization-focused evaluation*. Sage.
- Patton, M. Q. (2016). What is essential in developmental evaluation? On integrity, fidelity, adultery, abstinence, impotence, long-term Commitment, integrity, and sensitivity in implementing evaluation models. *American Journal of Evaluation*, 37(2), 250–265.
- Patton, M. Q. (2018). *Principles-focused evaluation: The GUIDE*. Sage.
- Patton, M. Q. (2019a, December 2). Climate emergency from a Blue Marble Evaluation perspective. *Blue Marble Evaluation Blog*. <https://bluemarbleeval.org/latest/climate-emergency-blue-marble-evaluation-perspective>
- Patton, M. Q. (2019b, December). Expanding futuring foresight through evaluative thinking. *World Futures Review*, 11(4), 296–307.
- Patton, M. Q. (2019c). Transformation to global sustainability: Implications for evaluation and evaluators. *New Directions for Evaluation*, 162(Summer), 103–118.
- Patton, M. Q. (2020a). *Blue marble evaluation: Premises and principles*. Guilford.
- Patton, M. Q. (2020b, March 23). Evaluation implications of the coronavirus global health pandemic emergency. *Blue Marble Evaluation Blog*. <https://bluemarbleeval.org/latest/evaluation-implications-coronavirus-global-health-pandemic-emergency>
- Patton, M. Q. (2020c). *Launch of Eval4Action*. <https://www.youtube.com/watch?v=SjsktJJQss&feature=youtu.be>
- Picciotto, R. (2009). Evaluating climate change and development. In R. D. van den Berg & O. Feinstein (Eds.), *Evaluating climate change, and development* (pp. 19–46). Transaction.

- Picciotto, R. (2013). The logic of development effectiveness: Is it time for the broader evaluation community to take notice? *Evaluation*, 19(2), 155–170.
- Picciotto, R. (2020a). From disenchantment to renewal. *Evaluation*, 26(1), 49–60.
- Picciotto, R. (2020b, January 23). The DAC criteria, personal email correspondence.
- Ramalingam, B., Singh, N. S., Mahieu, A., & Blanchet, K. (2020). Responding to covid-19. *Guidance for humanitarian agencies*. ALNAP.
- Richardson, R. (2018). Foreword. In S. Hussain & J. Vause (Eds.), *TEEB for agriculture & food: Background and objectives* (In TEEB for Agriculture & Food: Scientific and Economic Foundations) (p. 6). UN Environment.
- Rockström, J., & Sukhdev, P. (2016). *How food connects all the SDGs*. Stockholm Resilience Center. <https://www.stockholmresilience.org/research/research-news/2016-06-14-how-food-connects-all-the-sdgs.html>
- Rodin, J. (2014). *The resilience dividend: Being strong in a world where things go wrong*. The Rockefeller Foundation/Public Affairs.
- Rowe, A. (2019). Sustainability-ready evaluation: A call to action. *New Directions for Evaluation*, 162(Summer), 29–48.
- Rugg, D., Buehler, J., Renaud, M., Gilliam, A., Heitgerd, J., Westover, B., Wright-Deaguero, L., Bartholow, K., & Swanso, S. (1999). Evaluating HIV prevention: A framework for national, state and local levels. *American Journal of Evaluation*, 20(1), 35–56. <https://doi.org/10.1177/109821409902000104>
- Rugg, D. (2015, March 13). Key role of evaluation in new development agenda. *UN News*. <https://news.un.org/en/story/2015/03/493362>
- Rugg, D. (2016). The role of evaluation at the UN and in the new sustainable development goals: Towards the future we want. *Global Policy*, 7(3), 426–430. <https://doi.org/10.1111/1758-5899.12346>
- Schwab, K. (2017). We need a new narrative for globalization. *World Economic Forum*, 3/17. <https://www.weforum.org/agenda/2017/03/klaus-schwab-new-narrative-for-globalization/>
- Schwandt, T. (2018, January 9). DAC evaluation criteria. *Evaluation for Development Blog* comments. <http://zendaofircom/updates-dac-evaluation-criteria-part-5/>
- Schwandt, T. (2019). Post-normal evaluation? *Evaluation*, 25(3), 317–329
- Scriven, M. (2016). The last frontier of evaluation: Ethics. In S. I. Donaldson & R. Picciotto (Eds.). *Evaluation for an equitable society* (pp. 11–48). Information Age.
- Segone, M., & Tateossian, F. (2017). No one left behind. In M. Bamberegr, M. Segone, & F. Tateossian (Eds.) *Evaluation for Agenda 2030: Providing evidence on progress and sustainability* (pp. 23–34). <https://ideas-global.org/wp-content/uploads/2017/12/Chapter-2.pdf>
- Shanker, S. (2019a). *Definitional tension: the construction of race in and through evaluation*. Ph.D. Dissertation, University of Minnesota, Minneapolis.
- Shanker, V. (2019b). Tidying evaluation of “diversity” & “culture.” *AEA365 blog*, June 24. <https://aea365.org/blog/a-look-at-language-week-tidying-evaluation-of-diversity-culture-by-vidhya-shanker/>
- Slaper, T. F., & Hall, T. J. (2011). The triple bottom line: What is it and how does it work? *Indiana Business Review*, 86(1). <https://www.ibrc.indiana.edu/ibr/2011/spring/article2.html>
- Smith, L. T. (1999). Decolonizing methodologies: Research and indigenous peoples. Zed.
- Stame, N. (2018). Strengthening the ethical expertise of evaluators. *Evaluation*, 24(4), 438–451.
- Steel, T. (2018). *This is the ‘last generation’ that can save nature*. CNN News. <https://www.cnn.com/2018/10/29/health/wwf-wildlife-population-report-intl/index.html>
- Suarez, C. (2020, May 5). Sustainment: Moving beyond sustainability. *NPQ Blog*. [https://nonprofitquarterly.org/sustainment-moving-beyond-sustainability/?utm\\_source=NPQ+Newsletters&utm\\_campaign=8ad3ae2f2d-EMAIL\\_CAMPAIGN\\_2018\\_01\\_11\\_COPY\\_01&utm\\_medium=email&utm\\_term=0\\_94063a1d17-8ad3ae2f2d-12318013&mc\\_cid=8ad3ae2f2d&mc\\_eid=707ec28f50](https://nonprofitquarterly.org/sustainment-moving-beyond-sustainability/?utm_source=NPQ+Newsletters&utm_campaign=8ad3ae2f2d-EMAIL_CAMPAIGN_2018_01_11_COPY_01&utm_medium=email&utm_term=0_94063a1d17-8ad3ae2f2d-12318013&mc_cid=8ad3ae2f2d&mc_eid=707ec28f50)
- Systems in Evaluation. (2018). *Principles for effective use of systems thinking in evaluation*. Systems in Evaluation Topical Interest Group, American Evaluation Association. <https://www.systemsinevaluation.com/wp-content/uploads/2018/10/SETIG-Principles-FINAL-DRAFT-2018-9-9.pdf>

- Taleb, N. N. (2018). *Skin in the game: Hidden asymmetries in daily life*. Random House.
- TCC Group (2019). *Equity and evaluation: Models of how equity can and does impact evaluation*. <https://www.equitableeval.org/blog-main/2019/4/17/equity-and-evaluation-models-of-how-equity-can-and-does-impact-evaluation>
- Tolley, L. M. (2020, May 3). Learning, evaluating, and staying strong together. *AEA365 Blog*.
- Uitto, J. I. (2019). Sustainable development evaluation: understanding the nexus of natural and human systems. *New Directions for Evaluation*, 162, 49–68.
- United Nations Environment Program. (2019). The United Nations emissions gap report (released November 26). <https://www.unenvironment.org/resources/emissions-gap-report-2019>
- van den, Berg, R. D., & Feinstein, O. (Eds.). (2009). *Evaluating climate change, and development*. Transaction Publishers.
- van den, Berg, R. B., Magro, C., & Mulder, S. S. (Eds.). (2019). *Evaluation for transformational change: Opportunities and challenges for the sustainable development goals*. IDEAS. [https://ideas-global.org/wp-content/uploads/2019/11/2019-11-05-Final\\_IDEAS\\_EvaluationForTransformationalChange.pdf](https://ideas-global.org/wp-content/uploads/2019/11/2019-11-05-Final_IDEAS_EvaluationForTransformationalChange.pdf)
- Vataja, K., & Parkkonen, P. (2019, November 12). Evaluation as a reinforcer of democratic ideals: What is the role and purpose of evaluation in our post-normal era? Helsinki: The Finnish Innovation Fund Sitra. <https://www.sitra.fi/en/blogs/evaluation-as-a-reinforcer-of-democratic-ideals/>
- Vidueira, P. (2020, April 23). Aftermath of the 50th anniversary of the blue marble (earth) day. *Blue Marble Evaluation Blog*. <https://bluemarbleeval.org/latest/aftermath-50th-anniversary-blue-marble-earth-day>
- Westley, F. R., Tjornbo, O., Schultz, L., Olsson, P., Folke, C., Crona, B., & Bodin, Ö. (2013). A theory of transformative agency in linked social-ecological systems. *Ecology and Society*, 18(3), 27. <http://dx.doi.org/10.5751/ES-05072-180327>
- Westley, F. R., Zimmerman, B., & Patton, M. Q. (2006). *Getting to maybe: how the world is changed*. Random House Canada.
- Williams, B. (2018, January 7). DAC evaluation criteria. *Evaluation for Development Blog* comments. <http://zendaofircom/updates-dac-evaluation-criteria-part-5/>