

Developmental Evaluation in Theory versus Practice: Lessons from Three Developmental Evaluation Pilots

Heather L. Esper

William Davidson Institute at the University of Michigan

Yaquta K. Fatehi

William Davidson Institute at the University of Michigan

Rebecca Baylor

William Davidson Institute at the University of Michigan

Background: Developmental Evaluation (DE) practitioners turn to DE theory to make design and implementation decisions. However, DE practitioners can experience difficulty in fully understanding how to implement DE using theory because it is method agnostic (Patton, 2016). Instead, DE is a principle-based approach.

Purpose: This article presents an empirical examination of how DE theory was (or was not) applied during three DE pilots. Our analysis aims to better understand how DE theory is used in practice to expand the evidence base and strengthen future DE implementation.

Setting: A consortium of three organizations implemented three DE pilots through the United States Agency for International Development (USAID) from November 2016 to September 2019. The authors—who participated in the consortium—did not implement the DEs but instead conducted a study or meta-evaluation across the DE pilots.

Data Collection and Analysis: This article focuses on the results of an ex post facto analysis of three DE pilots based on the entire DE implementation experience. For each DE studied, we used mixed methods to collect data on the effectiveness of the DE approach, to identify adaptations to strengthen DE implementation in the USAID context, and to measure its value to stakeholders. Data included more than 100 hours of interviews, 465 pages of qualitative data, and 30 surveys completed by DE participants.

Findings: We find that the ability to apply the DE principles in practice is influenced, in no particular order, by DE participant buy-in to the DE, the Developmental Evaluator's aptitude, support and resources available to the Developmental Evaluator, and the number of DE participants. We also find that buy-in can change and this should be closely monitored throughout a DE to inform whether a DE should be paused or prematurely ended.

Keywords: *Developmental Evaluation; developmental evaluator skills; buy-in; DE practice; DE funder; meta-evaluation*

Background

In the field of evaluation, it is essential to bridge the gap between theory and practice. Research on how evaluation approaches and methodologies are written and implemented contributes to theories that are high quality, feasible to conduct in practice, and lead to valuable outcomes for evaluation end-users (Miller, 2010; Stufflebeam, 2001). Practitioners can help translate theory to practice. Learnings from practice can inform the adaptation of evaluation theory. This virtuous cycle ensures that the field of evaluation is continuously evolving and improving to meet the needs of evaluators and end-users better.

Our research seeks to better understand how DE theory is applied in practice. Examining the theory-practice relationship is particularly important and unique in the context of DE for a few reasons. First, DE is a methodologically agnostic approach that does not use prescribed steps (Patton, 1994, 2011, 2016). Instead, DE theory and, subsequently, its implementation are guided by eight essential principles: (1) *developmental purpose*, (2) *evaluation rigor*, (3) *utilization focus*, (4) *innovation niche*, (5) *complexity perspective*, (6) *systems thinking*, (7) *co-creation*, (8) and *timely feedback*. All eight principles should be applied and contextualized throughout the evaluation to ensure the evaluation's integrity (Patton, 2016).¹ However, existing case examples of DE demonstrate considerable variability in implementation in the field (Hayes et al., 2016; Patton et al., 2015). The flexible structure of DE warrants consideration because it impacts implementation decisions that affect whether the evaluation is true to the original ideals of how evaluators and users should carry out DE. Indeed, the Developmental Evaluator plays a pivotal role regarding what degree DE implementation aligns with DE theory. Second, experienced practitioners also report that it is common to find individuals claiming to be conducting a DE when they are not (Patton et al., 2015). At a minimum, it is worth

highlighting which aspects of DE implementation can benefit from improved clarity.

Essential Tenants of DE in Theory and Practice

DE is an evaluative approach aimed at facilitating the continuous adaptation of innovative interventions. It involves having one or more Developmental Evaluators integrated into the program team on a full-time or part-time basis to help guide adaptation to emergent and dynamic realities in complex environments (Patton, 1994, 2011, 2016). Developmental Evaluators adjust evaluation questions, methods, and analytic techniques using a range of evaluation methods as the project evolves to deliver emergent and contextualized utilization-focused findings on an ongoing basis. They participate in team meetings, document change, collect data, and share results with teams, all while remaining impartial. They also co-create recommendations and co-implement changes with DE participants (Baylor et al., 2020). A Developmental Evaluator is likely to build relationships and become an adviser to DE participants while continuing to maintain a third-party perspective. In such cases, they often become an active member of the teams' strategic adaptation and innovation. These features further make DE different from other forms of traditional evaluation.

The Developmental Evaluator captures, promotes, and uses emergent learnings during the DE to inform DE participants and their resulting actions. Emergent learnings are defined as new programmatic or environmental developments, such as new information gained or changes in existing stakeholder relationships. The Developmental Evaluator shares this data with participants and uses findings and insights to inform adjustments to the program's *intended* strategy.² Intended strategies are developed during the program design phase. Adjustments to the program result in the emergence of *realized* strategies to achieve the

¹ We define DE theory based on Michael Quinn Patton's work. As a result, the eight DE principles are a foundational component of our assessment.

² DEs can be implemented at any level: strategy, program, activity, etc.

program's goals (Mintzberg, 2007). At the same time, because of the Developmental Evaluator's findings, other previously intended activities are never implemented (i.e., they are *unrealized* strategies of the original program design).

At the heart of DE theory and practice are eight essential principles (see Table 3). However, there is no step-by-step guide or checklist for when or how to embed the principles throughout the day-to-day execution of the approach. Patton (2016) suggests DE practitioners promote flexibility and focus on the degree of manifest sensitivity³ to each of the essential principles during the evaluation. To implement DE, therefore, evaluation practitioners should understand each principle and how it applies to the context. They must also make intentional efforts to apply the principles. However, the degree to which they are used is subject to the Developmental Evaluator's discretion based on the needs of key users.

Without understanding, thinking about, and attempting to apply each principle, it is impossible to foster the "DE mindset" needed to guide implementation (Patton, 2016). While Patton (2016) and other DE practitioners (Beer, 2019; Beer et al., 2019; Blanchet-Cohen & Langlois, 2010; Gamble, 2008; Lam 2016; Lam & Shulha, 2015; Langlois et al., 2012; Patton et al., 2015) have shared vital considerations about how to design and conduct the DE approach, existing literature does not go so far as to systematically evaluate how DE theory is applied in practice.

Examining DE Theory in Practice

We examine the relationship between DE theory and practice using three criteria: (1) *operational specificity*, (2) *range of application*, (3) and *feasibility of practice* (Miller, 2010).⁴

³ Manifest sensitivity refers to the degree to which each principle is embodied in the DE design and use of findings. Whereas other evaluation theories may be "operationalized," the term "manifest sensitivity" is used because the DE principles serve as sensitizing concepts that are not comprised of a set of step-by-step rules that can be followed.

⁴ The analysis focuses on challenges that evaluation practitioners may experience when they implement DE by applying the eight essential

Operational specificity examines the clarity of guidance associated with an evaluation theory, including implications for practice. Since DE is a principles-based approach that is not operationalized with a standard set of activities, we use the operational specificity criterion to focus on the degree of manifest sensitivity achieved for each principle. For simplicity's sake, we use the terms "operationalization" and "manifest sensitivity" interchangeably in our analysis. The *range of application* considers the limitations of the application of the evaluation theory, and *feasibility in practice* explores how easily the theory can be applied in practice. We chose not to include two criteria—(1) *discernable impact* and (2) *reproducibility*—because of their lack of relevance to DE.⁵

Together, the three criteria provide lenses to examine how DE theory is applied in practice and the circumstances that allow successful DE to occur. We use these criteria to examine if and how the DE principles were applied in each pilot. Where relevant, we also examine the reasons why these principles could not be adequately applied. Applying these three criteria allows us to understand the application of DE theory in practice, as well as identify foundational circumstances to DE success. Thus we use these three criteria to explore the following question: *How can we strengthen the DE approach in practice?*

principles. It is not an examination of the quality of DE as a theory.

⁵ We conducted a phone interview with Robin L. Miller, who developed the framework used to examine the relationship between evaluation theory and practice, to understand the frameworks potential to apply it to DE. During the call, we learned it is not necessary to use all five criteria for such an assessment.

Research Context and Study Design

Sample

Since 2016, a consortium of three organizations called the Developmental Evaluation Pilot Activity (DEPA-MERL)⁶ has conducted three DE pilots for USAID (see Table 1). While we the authors of this article are part of the consortium we did not implement any pilots and instead played a research role in the consortium. Similar to the experience of other DE implementers (Hayes et al., 2016), the DEPA-MERL Consortium initially had limited familiarity with DE implementation. The duration of the DE pilots spanned from 9 to 22 months. Each DE included the following:

- One full-time Developmental Evaluator physically co-located with the DE participants;
- A DE Administrator (a person or group of people) responsible for (a) launching and overseeing the DE, (b) providing technical support to the Developmental Evaluator on an as-needed basis, and (c) liaising with the funder;
- A DE Point of Contact from USAID, either a Contracting Officer's Representative or a DE Activity Manager from the DE participants' team.

To select these pilots, DEPA-MERL vetted several potential DE opportunities, of which only three moved into the DE start-up phase. To vet a potential DE opportunity, DEPA-MERL first assessed whether the programmatic conditions were favorable for DE and whether another evaluation methodology might be more appropriate (DEPA-MERL Consortium, 2019a). DEPA-MERL also held a series of initial conversations with key stakeholders from the related program to assess their interest in and

readiness for DE, as well as to understand whether DE is a good fit for the program's context and needs (DEPA-MERL Consortium, 2019a). Next, DEPA-MERL developed a preliminary scope of work (SOW) for the DE, including the rationale for DE, purpose and use of the DE, and potential/illustrative evaluation questions. The SOW was then circulated to potential DE stakeholders for their review and comments.

The finalized SOW was used to identify the unique skills needed for each pilot's Developmental Evaluator. DEPA-MERL sought Developmental Evaluators with a diverse analytical toolbox, strong interpersonal skills, the ability to communicate effectively with different types of stakeholders, and experience facilitating organizational change. DEPA-MERL screened and interviewed all candidates. Key program stakeholders also interviewed top candidates to assess fit with their team.

Upon selection, the DEPA-MERL worked with USAID stakeholders to rapidly integrate the Developmental Evaluator into the team. The Developmental Evaluator received relevant background materials and a list of key contacts and stakeholders to meet. DEPA-MERL also worked to ensure the inclusion of the Developmental Evaluator in relevant program communications and meetings. Within a month or two of the DE starting, the Developmental Evaluator held an Acculturation Workshop with stakeholders to kick off the DE. The workshop provided an overview of DE, explained the role of the Developmental Evaluator, and allowed stakeholders an opportunity to refine the evaluation questions and develop a draft work plan. The workshop also allowed stakeholders to establish shared expectations and communication protocols.

⁶ The U.S. Global Development Lab's Monitoring, Evaluation, Research, and Learning Innovations Program at USAID commissioned DEPA-MERL.

Table 1
Description of Three DEPA-MERL DE Pilots

Name (nickname)	<i>Supporting collective impact in Cambodia (Partner Relations DE)</i>	<i>Advancing uptake of innovations at USAID (Innovations DE)</i>	<i>Examining knowledge management practices (Knowledge Management DE)</i>
Dates	November 2016–March 2018	March 2017–December 2018	December 2018–September 2019
Project context	A network of nearly 60 organizations working to increase the number of children living in safe and nurturing family-based care in Cambodia	Innovation hub at USAID headquarters in Washington, DC	Large USAID bureau in Washington, DC
Rationale for choosing DE	The direction of certain aspects of the initiative was uncertain, not all parts of the program were clearly defined; the project was complex owing to the large number of partners involved, and the initiative had a focus on iterative piloting, learning, and growing cycles of innovation. Lead DE organizations, including USAID, asked for DE to support building cross-sector alignment and a culture of trust and respect across partners	Hub had a history of applying learnings to evolve programming; however, learning had not been approached systematically	Bureau was undergoing a complex transformation, and DE was viewed by senior leadership as an approach that could help inform knowledge management strategies
DE SOW	DE focused on understanding how partner relations impacted the work. Explored what was and wasn't working regarding management, coordination, and collaboration across partners	The six lab teams and offices conducted the DE. DE assessed and provided learnings about the sustained uptake of innovations promoted by USAID, including how it differed in various USAID sub contexts. The DE also helped the teams develop their exit strategies Note: The SOW and time line of the DE expanded twice. The expansions included adding (1) additional teams to the DE, which increased the scope and time line, and (2) time to allow for sustainability planning and capacity building	The goal was to help the bureau develop a knowledge management process with five teams that would transition over to a proposed new bureau created under the USAID Transformation process. However, the attempted DE did not advance beyond the start-up phase, as no agreement was reached on a defined SOW for the DE

Name (nickname)	Supporting collective impact in Cambodia (Partner Relations DE)	Advancing uptake of innovations at USAID (Innovations DE)	Examining knowledge management practices (Knowledge Management DE)
Evaluator skills	<p>All three Evaluators had the following skills: extensive technical knowledge; skills and expertise in evaluation design, concepts, and approaches; strong analytical skills to support both quantitative and qualitative analysis; excellent oral and written communication and report writing; keen attention to detail; leadership and strategic thinking skills; facilitation skills; respect for local stakeholders and partners; experience working for multiple stakeholders with competing priorities; ethical research (maintain confidentiality and anonymity when required); flexibility; energy; humility; willingness to learn on the go; ability to resolve conflicts; and active listening</p> <p>Programming and evaluation experience in child protection, child rights, education, or related sector; experience in international development; and experience evaluating complex systems</p>	Experience in scaling in international development and familiarity with USAID context	Experience working on developing knowledge management systems and familiarity with USAID context
Example of types of methods used by Evaluator during DE	Document reviews, observations, team meetings participation, a formal literature review, key informant interviews (KII), terms of reference sessions, an assessment rubric, a partner survey, and case studies	Focus groups, KIIs, observations, team meetings participation, document review, facilitated workshops, partner maps, case studies, qualitative coding, process tracing, positive deviance case studies, and outcome harvesting	Observations, team meetings participation, and KIIs
Embeddedness	Embedded full time in partner's office in Phnom Penh	Embedded full time in USAID's DC office	Full-time Evaluator never fully embedded for several reasons, including the government shutdown in early 2019, office renovations that required all staff to work from home, and no formal decision on which subteam to be embedded within
Links to final reports	Identifiable information removed for the review process	Identifiable information removed for the review process	Identifiable information removed for the review process

The first two pilots were successful DEs, but the third DE pilot, 'Knowledge Management DE', did not evolve past the DE start-up phase. The Developmental Evaluator was not able to incorporate all the DE principles. We consider a successful DE to be one in which DE captured, promoted, and enabled utilization of emergent learnings gained through the DE. Thus the 'Knowledge Management DE'

provided a learning opportunity to examine why the pilot did not work in practice (Baylor et al., 2019).

Data Collection and Analysis

We identified outcomes of the DE using the outcome harvesting⁷ methodology in the two successful pilots. We selected this methodology because it allowed us to examine how individual outcomes contributed to system-wide changes (Gold et al., 2014). We used outcome harvesting to assess any outcomes (behavior, relationship, action, policy, and/or practice of stakeholders)⁸ that the DE contributed to, either directly or indirectly. Contributions included the Developmental Evaluator providing documentation, collecting data, developing recommendations, and promoting ideas and best practices.

The primary data sources for harvesting the outcomes included (1) an event log completed monthly by the Developmental Evaluator to track emergent learnings and (2) a monthly interview that the authors conducted with the Developmental Evaluator to explore priority emergent learnings. Descriptions of the outcomes included the Developmental Evaluator's contribution to the outcome and the significance of the outcome to the program. DEPA-MERL and the Developmental Evaluator reviewed these outcome descriptions quarterly to update them with any new information. At the end of each DE, the authors held multiple meetings with the Developmental Evaluator to edit and finalize all outcome descriptions before substantiation.

The authors selected outcomes for substantiation to ensure that (1) they included outcomes where the DE captured, promoted, or enabled the utilization of key emergent learnings within the initiative and (2) by verifying this subset of outcomes, the remaining harvested outcomes (e.g., those that were not being substantiated) would automatically be validated or indirectly substantiated because of the

interconnectedness of the outcomes within a theme.⁹ Semi-structured interviews were conducted with substantiators to verify the accuracy of outcomes.

After substantiation interviews, we determined the substantiators' level of agreement with each outcome description and contribution of the DE to the outcome. The authors also incorporated new data captured during the substantiation interviews. DEPA-MERL then reviewed and reverified each updated outcome for accuracy.

We categorized each of these outcomes to understand the types of changes that resulted from the DEs (see Table 2; Fatehi et al., 2018):

1. **Role of the DE.** Did the DE *capture*, *promote*, or *enable* the utilization of emergent learning(s) associated with the outcome?
2. **Type of Change.** Did a particular outcome reflect changes primarily related to the *knowledge* and *capabilities* of stakeholders? Their *engagement* and *relationships*? Or did more formal *institutional* and *policy* changes occur in the program? To the strategy used by a participating team?
3. **Orientation of Change.** Did a particular outcome have *positive*, *negative*, or *both positive* and *negative* effects on the program in the short term?

⁷ In the outcome harvesting method, researchers "collect (harvest) evidence of what has changed (outcomes) [in the program] and then, working backwards, determine whether and how an intervention has contributed to these changes" (Wilson-Grau, 2015).

⁸ In this study, we expanded the definition of what constitutes as an outcome to also include products

and deliverables produced by DE to answer the research question.

⁹ We consulted with Ricardo Wilson-Grau, the cofounder of the outcome harvesting approach on how to strengthen the proposed substantiation process given the unique challenges of DE.

Table 2
Types of DE Outcomes in the Two Successful DEPA-MERL Pilots

		Role of the DE*			Type of change (for the program/DE participants)				Impact of change		
		Capture	Promote	Enable utilization	Engagement and relationship	Institutional and policy	Strategy	Knowledge and capability	Positive	Negative	Both
Supporting collective impact in Cambodia	# outcomes	3	3	11	6	7	N/A	4	11	2	4
	% of all outcomes	18	18	65	35	41	N/A	24	65	12	24
Advancing uptake of innovations at USAID	# outcomes	3	5	14	3	2	8	9	17	0	5
	% of all outcomes	14	23	63	14	9	36	41	77	0	23

*Please note that any harvested outcomes categorized as *enabled utilization* of emergent learnings are outcomes the Developmental Evaluator *captured* and *promoted* first. However, to avoid double-counting, each harvested outcome was only categorized with one code (i.e., “enable utilization”).

In addition to outcome harvesting, the authors explored barriers and enablers to the success of the DE. During the monthly reflection interview with the Developmental Evaluator and in the substantiation interviews with stakeholders, the authors asked open-ended questions to learn of barriers and enablers to the implementation of the DE. These included factors that interviewees experienced or faced and were particular to the program, sector, and/or local context. Using NVivo, the authors conducted line-by-line coding for barriers and enablers using summaries from these interviews. The team ensured intercoder reliability of 80%–90% on all codes or factors and discussed any coding-related discrepancies during internal weekly meetings. The team developed codes through both deductive and inductive processes.

To answer our research question, *How can we strengthen the DE approach in practice?*, we developed a table (see Table 3) to assess the

degree of manifest sensitivity (i.e., operational specificity) achieved for each principle in each pilot (we used a scale of *yes*, *partially*, and *no*). We did this because DE is a principle-based approach. To develop the table, the authors conducted an in-depth review of the DE pilot reports, harvested outcomes, and related interview data to determine the extent to which the DE principle was applied. A representative from each partner in DEPA-MERL, as well as one of the Developmental Evaluators, reviewed the table. The authors adjusted the rating of the application of the principles accordingly based on these reviews. Using the table, we then assessed what circumstances allowed or didn't allow the application of principles in each pilot. We explore the factors that limit the application of DE theory in the “Range of Application” section, and factors that influence how easily DE theory can be applied in practice are explored in the “Feasibility in Practice” section.

Findings and Discussion

Operational Specificity/Manifest Sensitivity

Since DE is a principles-based approach that is not operationalized with a standard set of activities, we use the operational specificity criterion to focus on the degree of manifest sensitivity achieved for each principle.

The Developmental Evaluator experienced various levels of difficulty incorporating each of the principles into the three DE pilots. The “Innovations DE” had the most principles incorporated, whereas the “Knowledge Management DE” had the fewest (see Table 3). To categorize the principles based on whether they were “easy” or “hard” to operationalize, we used the following scoring system based on the results in Table 3: yes = 1 point; no = 0 points; partially = 0.5 points. The easiest principles to operationalize at 2.5 points each were *developmental purpose*, *evaluation rigor*, *complexity perspective*, and *systems thinking*. Meanwhile, the most difficult principles to operationalize at 2 points each were *innovation niche*, *utilization focus*, *co-creation*, and *timely feedback*. Even though the Developmental Evaluator influenced the degree to which principles were applied, we note that it is worth exploring using resources to share these principles with DE funders and critical participants to ensure that they are consistently applied.

We explore factors that influence the degree of manifest sensitivity for each principle in the forthcoming sections that focus on the *range of application* and *feasibility in practice* criteria. These factors include the DE participants’ (including funders) buy-in, the Developmental Evaluator’s aptitude, the support and resources available to the Developmental Evaluator, and, lastly, the number of DE participants.

Range of Application

DE is not appropriate for every situation. To bridge the theory-practice gap and ensure fit, DE practitioners should investigate the context¹⁰ under which the DE will be implemented and which contingencies will govern and limit the application of DE. Literature states that DE should be used as an evaluation approach when the program is operating in a complex environment where nonlinear change occurs or when the program’s theory of change is unknown (Patton, 2011) or when both occur.¹¹ We determine that all three USAID pilots met this essential criterion. We also find that DE participant buy-in is necessary at the start of a DE, as well as critical throughout the DE when examining the manifest sensitivity for each DE principle across the pilots.

Based on our study, we find that most successful DE pilots occur in organizations and programs where the DE participants’ culture reflects comfort with adaptive work (i.e., learning while doing and experimentation; Gamble, 2008). In our pilots, successful DEs had DE participants (including funders) who from the start of the DE were active recipients of the DE process. They enthusiastically engaged and committed to co-creating a DE that exhibited the ideals of rigorous evaluative thinking, a systems perspective, and timely feedback. Theory deems that Developmental Evaluators can make recommendations but not the decisions to execute them, and hence active engagement by DE participants is significant for the success of the DE. In the pilots, DE participants’ reflected buy-in through conscious efforts to integrate the Developmental Evaluators into their teams.

¹⁰ For more information on what contexts are appropriate to support DE, see Chapter 1 of Patton et al. (2015).

¹¹ These are especially cases where traditional evaluation methods cannot measure impact.

Table 3
Eight Essential DE Principles Applied in Practice

Definitions—DE principles (Patton, 2016)	What to look for to assess the degree of manifest sensitivity in DE practice (Patton, 2016)	Was the essential DE principle applied in practice?		
		Partner Relations DE	Innovations DE	Knowledge Management DE (attempted DE)
<p>1. Developmental Purpose Principle. Illuminate, inform, and support what is being developed by identifying the nature and patterns of development (innovation, adaptation, and systems change) and the implications and consequences of those patterns</p>	<p>The innovation is being developed. The evaluation captures what is being developed and associated implications. The evaluation itself is developed as the innovation develops</p>	<p>Yes. The Developmental Evaluator and DE participants co-identified the purpose of the DE, which indeed had an aspect of innovation. All key participants showcased their buy-in and readiness. The DE evolved based on the priority needs of the DE participants</p>		<p>Partially. The purpose of the DE, while identified, did not get buy-in from all key DE participants. The Developmental Evaluator faced many challenges, such as high staff turnover, limited access to DE participant’s time and data. Some of the challenges the Developmental Evaluator faced were because the DE pilot coincided with two external events: the longest government shutdown in U.S. history and the USAID Transformation process</p>
<p>2. Evaluation Rigor Principle. Ask probing evaluation questions; think and engage evaluatively; question assumptions; apply evaluation logic; use appropriate methods; and stay empirically grounded— that is, rigorously gather, interpret, and report data</p>	<p>DE is empirically driven, and evaluative thinking is the foundation of all aspects of the engagement</p>	<p>Yes. The Developmental Evaluator conducted rigorous evaluation activities, from KIIs and administering surveys with DE participants, to analyzing data using qualitative and quantitative analyses. The Developmental Evaluator also conducted workshops to share findings, engage participants in evaluative thinking, and hear their interpretation of the data and insights</p>		<p>Partially. When provided access, the Developmental Evaluator gathered and interpreted data to the best of her ability given her poor integration into the teams. DE participants did not always show interest in evaluative thinking, the collected data, or sharing feedback with the Developmental Evaluator</p>
<p>3. Utilization-Focused Principle. Focus on the intended use by end-users from beginning to end, facilitating the evaluation process to ensure utility and actual use</p>	<p>DE data is collected with the intended use by end-users in mind</p>	<p>Yes. While decision making took time, DE data was used by the intended DE participants to develop future strategies and adapt the</p>	<p>Yes. DE participants used DE data to develop future strategies and adapt their innovation. The Developmental Evaluator</p>	<p>No. DE data was not used by participants to illuminate and adapt strategies. The Developmental Evaluator was viewed by DE participants as an “extra set of hands” to complete one-off evaluative</p>

Definitions—DE principles (Patton, 2016)	What to look for to assess the degree of manifest sensitivity in DE practice (Patton, 2016)	Was the essential DE principle applied in practice?		
		Partner Relations DE	Innovations DE	Knowledge Management DE (attempted DE)
		innovation. Several utilization-focused pieces of work moved forward, including large demand-driven events to support learning uptake	conducted activities and facilitated meetings with DE participants to use findings	activities. No process was in place to ensure DE utility and actual use of findings. No decision maker was present to enable use
4. Innovation Niche Principle. Elucidate how the change processes and results being evaluated involve innovation and adaptation, the niche of DE	Agreement to innovate is explicit and authentic	Yes. DE participants were unhappy with the status quo of their program and worked with the Developmental Evaluator to innovate their organizational structures and social relationships. Innovations also took place in program strategy and processes	No. The Developmental Evaluator was unable to help develop or adapt the innovation within the program. There was no widespread engagement or commitment to innovate by DE participants	
5. Complexity Perspective Principle. Understand and interpret development through the lens of complexity and conduct the evaluation accordingly. This means using complexity premises and dynamics to make sense of the problems being addressed; to guide innovation, adaptation, and systems change strategies; to interpret what is developed; to adapt the evaluation design as needed; and to analyze emergent findings	The complexity of the system in which the innovation and evaluation are occurring is described. The complexity of the innovation being developed and evaluated is also described. The DE design, process, and outcomes take into account the identified complexity	Yes. The Developmental Evaluator acknowledged the high contextual complexity, both internal and external to the project. The Developmental Evaluator worked with teams to understand complex dynamics, as well as to adapt their innovation and strategy to account for the high level of complexity they faced	Partially. While the Developmental Evaluator and DE participants recognized the complexities affecting the program, the DE participants lacked sufficient coherent organizational context to support the DE. The program, the teams, and the overarching bureau were undergoing a restructure, which led to confusion, low transparency in data, and an inability to guide adaptation	

Definitions—DE principles (Patton, 2016)	What to look for to assess the degree of manifest sensitivity in DE practice (Patton, 2016)	Was the essential DE principle applied in practice?		
		Partner Relations DE	Innovations DE	Knowledge Management DE (attempted DE)
<p>6. Systems Thinking Principle. Think systematically throughout, being attentive to interrelationships, perspectives, boundaries, and other key aspects of the social system and context within which the innovation is being developed and the evaluation is being conducted</p>	<p>Interrelationships, perspectives, and boundaries inform the innovation and DE</p>	<p>Yes. The Developmental Evaluator was very attentive to relationships and perspectives, as the program had a large number of diverse organizations (size, culture, language, mission). The DE also focused on governance structures</p>	<p>Yes. The Developmental Evaluator facilitated connections among DE participants and collaborated with them to understand critical aspects of the social system and context</p>	<p>Partially. The Developmental Evaluator could not get sufficient access to key DE participants to develop a complete understanding of the perspectives, boundaries, and interrelationships among DE participants</p>
<p>7. Co-creation Principle. Develop the innovation and evaluation together—interwoven, interdependent, iterative, and co-created—such that the DE becomes part of the change process</p>	<p>A collaborative process that is active, reactive, interactive, and adaptive emerges from a mutually trusting relationship between the Development Evaluator and the DE participants</p>	<p>Partially. The Developmental Evaluator acknowledged the need to build mutually trusting relationships with DE participants to facilitate an effective and collaborative DE process. DE data that would influence the direction of the DE was not always shared with DE participants first and instead needed to be cleared by USAID</p>	<p>Yes. The Developmental Evaluator worked with DE participants to develop the innovation and the evaluation together. The DE created space for participants to pause and reflect on data, and the DE became part of the change process</p>	<p>Partially. The Developmental Evaluator worked hard to coordinate an Acculturation Workshop where the DE's initial evaluation questions were co-created. A few vital persons could not attend this event, and it is not fully understood if their absence affected the success of the workshop. However, the Developmental Evaluator was never fully integrated into the DE participant teams, preventing the DE from becoming an interwoven, co-created part of the program's change process</p>

Definitions—DE principles (Patton, 2016)	What to look for to assess the degree of manifest sensitivity in DE practice (Patton, 2016)	Was the essential DE principle applied in practice?		
		Partner Relations DE	Innovations DE	Knowledge Management DE (attempted DE)
<p>8. Timely Feedback Principle. Timely feedback to inform ongoing adaptation as needs, findings, and insights emerge rather than only at predetermined times (e.g., quarterly or at midterm and the end of the project)</p>	Data gathered by the DE and feedback of DE findings are timely and ongoing	<p>Partially. The Developmental Evaluator worked closely with some DE participants to share data regularly. However, in some cases, USAID wanted to see DE data first, which slowed down the Developmental Evaluator's ability to provide timely feedback</p>	<p>Yes. The Developmental Evaluator met and shared data with DE participants every week. Data was shared and reflected upon as needed and in as close to real time as possible. A large majority of participants said the DE was much better at providing timely feedback than other evaluation approaches¹</p>	<p>Partially. When requested, the Developmental Evaluator met and shared data with DE participants promptly. However, DE participants were resistant or unable to share data about their program openly and transparently. Interaction with DE participants was limited and often occurred via email</p>

¹ In a survey that was conducted with participants of the second DE pilot, 73% (11 of 15 respondents) said their DE was *much better* than other more traditional evaluation approaches because DE allowed for awareness of complexities of the local environment and was timely in providing feedback.

The more embedded the Developmental Evaluators were in the team, the more they could capture, promote, and use key data to support adaptations. Hence, we determine that both continuous buy-in from DE participants to engage in continuous learning and adaptation and conscious integration of the Developmental Evaluator are necessary conditions to apply DE successfully.

Further, as the literature states, DE serves stakeholders well during times of high uncertainty and complexity (Patton, 2011). However, it may be possible to have “too much” ambiguity surrounding a program situated in a complex setting such that it can be challenging to apply the DE principles effectively. During a DEPA-MERL DE clinic, Michael Quinn Patton, considered the founder of DE, shared that “DEs require a coherent organizational context to be effective.” DEs perform best when there are well-defined goals and objectives and a clear administrative structure (Patton, 2019). Although the Knowledge Management DE pilot started with strong buy-in, the high uncertainty stemming from the USAID Transformation process was the main impediment to making the DE a success. The Transformation process led to uncertainty about the future direction of the DE and the program. A high level of turnover among intended DE participants, including leadership, further exacerbated the issue. DE participants and decision makers had competing priorities and limited capacity to engage in the DE. As a result, participant buy-in dropped. We found that buy-in needs to be carefully monitored throughout the DE. While there certainly may be situations where a DE launches with strong buy-in, the DE can later be deemed unfeasible to continue because of waning support (such as through leadership turnover or an external event halting DE progress). If a DE has failed to gain sufficient momentum with stakeholders, then the DE process should be ended prematurely to safeguard resources and time.

To summarize, strong buy-in is needed at the start of the DE among key DE participants, but support can waver throughout the DE

and, hence, needs to be continually monitored. If buy-in decreases, decision makers should take steps to build support or assess whether the decreased support is threatening the ability of the DE to progress. Before starting a DE, decision makers should ask themselves, Do DE participants (including funders) have a culture of learning while doing and experimenting such that they will actively engage in the DE, as well as consciously integrate the Developmental Evaluator into the program? These steps can strengthen the DE approach in practice.

Feasibility in Practice

Assuming that context and stakeholder buy-in (including the integration of the Developmental Evaluator) are deemed suitable, the ability to implement DE depends on the Developmental Evaluator’s technical and interpersonal skills, the additional support and resources available to the Developmental Evaluator, and the number of DE participants.

One of the hardest aspects of starting a DE is finding a suitable Developmental Evaluator. The challenge arises because decision makers are looking for someone who has the necessary “hard” technical monitoring and evaluation skills, as well as the “soft” interpersonal skills required to conduct activities with DE participants. These skills range widely from the ability to speak “truth to power,” track the development process, listen and communicate, tolerate ambiguity, and engage DE participants in process facilitation and sense-making (Baldwin & Lander, 2018; Blanchet-Cohen & Langlois, 2010; Gamble, 2008; Hayes et al., 2016; Patton et al., 2015). During the pilots, the Developmental Evaluators conducted evaluative and adaptive activities as part of their embedded work with DE participant teams (DEPA-MERL Consortium, 2019a, 2019b). Evaluative activities included data collection and analysis to develop insights and answer key research questions of the DE.¹² Adaptive activities included facilitation and organizational

adaptation given the longevity of a DE, the timeliness of DE feedback, and the integrated relationship of the Developmental Evaluator.

¹² Activities may include interviews, focus groups, or other methods similar to those conducted in other types of evaluations. However, one key difference is that the activities may require

change support to enable DE participants to implement adaptations. Their work extended beyond sharing emergent learnings and recommendations with the DE participants to helping facilitate action, change, and adaptation. The support (or lack thereof) that Developmental Evaluators receive can affect their ability to perform all these activities. In addition, Developmental Evaluators use a range of strategies to generate buy-in and successfully embed into teams, including conducting “quick-win” tasks early in the DE. These tasks allow Developmental Evaluators to demonstrate their skills and value to the DE participants and may not be directly related to the DE’s line of inquiry. Because so much was expected of the Developmental Evaluators throughout all the pilots, DEPA-MERL learned that providing support to help manage their workloads was a beneficial use of resources. This support makes DE more feasible in practice. As one of the Developmental Evaluators shared,

It’s great to have a [supporting] team member with some dedicated time to bounce ideas off of, to talk through what types of tools or approaches you’re thinking of taking, and most importantly...someone to pull you [the Developmental Evaluator] out of the weeds once a week or so and help you see the bigger picture, re-orient to the DE research question, remind you when things are getting personal, etcetera.

In DEPA-MERL, the DE Administrator played a valuable role in ensuring that the Developmental Evaluator had the support and resources necessary to conduct rigorous evaluative and adaptive activities simultaneously. The DE Administrator helped the Developmental Evaluator problem solve, assisted him or her in seeing the bigger picture by taking a moment to zoom out of the details, encouraged the Developmental Evaluator to maintain objectivity with DE participants, and provided extra support to carry out technical work as needed. Also, the DE Administrator assisted the Developmental Evaluator in hosting an Acculturation Workshop. The workshop itself was an essential step in enabling the feasibility of practice. The workshop enabled a shared understanding of how the Developmental Evaluator’s DE work would unfold from the outset.

The number of DE participants can also influence the feasibility of the DE. The DE in Cambodia was of a program of nearly 60 organizations working together as an extensive network. The large number of teams participating in the DE required a large amount of the Developmental Evaluator’s time to manage stakeholder relationships. It also made it challenging to create utilization-focused outputs that met all participant’s needs. Unequal engagement of the DE participant teams led to problems applying the utilization-focus principle. The authors find that these three variables affect the DE approach in practice.

Conclusions: Developing Stronger Theory Using Evidence From Three DE Pilots

In the open and ambiguous world of DE, “failing forward”—learning from the mistakes that we make—is essential (Hayes et al., 2016). Developmental Evaluators, practitioners, and researchers should document lessons and study DE implementation to continue to use and strengthen this approach in practice. Indeed, a lack of specification about how to operationalize DE theory in practice has contributed to individuals and organizations calling their implementation a DE when it is not. Such miscommunication further adds confusion to what is and is not a DE.

While we understand that the aforementioned findings were learned in the USAID context, the authors believe that these lessons can be applied universally. We also recognize that our research is composed of a small sample of DEs (three) and that there are stark differences between them, such as project context, the SOW, and varied skills of the Developmental Evaluators. Despite the small sample size and differences, we are confident in our findings. This is because the DE Administrator was the same across the pilots and provided common support in all three pilots. This afforded a unique opportunity to compare and contrast across DEs.

We find that some of the DE principles have a higher degree of manifest sensitivity

than others. Although it is primarily the responsibility of the Developmental Evaluator to ensure manifest sensitivity to the principles, consider making key DE participants (including the funder) aware of the DE principles during implementation. They can play a role in helping apply principles that tend to be more difficult to operationalize. They can also help ensure that DE principles are consistently applied throughout the evaluation from DE design, data collection, feedback, and adaptation of the innovation. To conclude, to strengthen the DE practice, first, the Developmental Evaluator should consciously embed the eight DE principles throughout the DE. Second, the Developmental Evaluator should consider using resources to inform the DE stakeholders and funders of the DE principles.

Most notably, our findings demonstrate that there is a range of application of DE. Current literature states that DE cannot be selected as the approach when there is not enough buy-in from critical stakeholders. However, our findings also demonstrate that buy-in should be monitored throughout the DE to strengthen the DE approach in practice. When buy-in decreases during the DE to a critical level because of leadership turnover, incredible complexity, or competing priorities, the Developmental Evaluator should assess whether it is appropriate to continue the DE.

Our findings suggest several variables influence the feasibility of successfully implementing the DE approach. DE requires a great deal of effort from a Developmental Evaluator to be successful. Having a Developmental Evaluator with excellent technical and exceptional interpersonal skills is critical to the success of a DE. It is helpful to have a DE Administrator to provide additional managerial and technical support to the Developmental Evaluator to increase capacity given the role Developmental Evaluators must fill. The number of DE participants can also influence the feasibility of implementing DE, in particular the utilization-focused principle.

To conclude, our analysis revealed key variables that should be considered at different points in time to strengthen the DE approach in practice. First, DE participant buy-in is a critical variable in deciding whether the DE methodology should be used or not.

Second, variables to consider when designing a feasible DE include the Developmental Evaluator's aptitude, support, and available resources, as well as the number of DE participants. The Developmental Evaluator and DE Administrator should monitor these variables using pulse-check reflection sessions, brief surveys, or targeted KIIs to course correct or pause the DE if needed.

Areas for Future Research and Guidance in DE Literature

To help inform guidance related to DE's operational specificity, range of application, and feasibility in practice, we considered the manifest sensitivity to the principles and the strategic use of a Developmental Evaluator during DE implementation. However, other areas may influence DE effectiveness and would benefit from increased guidance in literature, including the following:

- **DE Team Structure.** Either one or more part-time evaluators, or one or more full-time evaluators, or a combination of both can conduct a DE. Further, a DE Administrator or other technical support team may support Developmental Evaluators, but the benefit and structure of such support teams are unknown.
- **DE Funding Structure.** In theory, all DE participants should have an equal say in the DE. However, in practice, it is common to see a large institution fund a DE, which can influence buy-in. Often, DE funders participate in the DE, which can influence how and when findings are shared and which adaptations are approved. In addition, multiple funders may contract a DE. The funders' relationship to the DE may vary and require clarification at the outset of the evaluation. For example, one funder may also be a direct participant and recipient of the DE findings and processes.
- **Relationship Between the Developmental Evaluator and the DE Participants.** In theory, an external person selected through a competitive process should

conduct DEs. In practice, there is often interest in conducting an internal DE by an in-house person or team for the sake of resources or ease of finding the Developmental Evaluator.

Decisions related to the above areas have the potential to impact the success of the DE and the ability of DE theory to be applied in practice.

Acknowledgments/Funding

This research was made possible by the support of the American people through USAID. The contents of this article are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the US government. The authors would like to thank reviewers who helped shape and strengthen the paper (listed in alphabetical order): Danielle de Garcia, Rebecca Herrington, Robin Miller, Michael Quinn Patton, and David Yamron.

References

- Baldwin, C. K., & Lander, R. (2018). Developmental evaluator functional role activities and programmatic developments: A case study analysis. *American Journal of Evaluation*, 40(1), 35–54. <https://journals.sagepub.com/doi/10.1177/1098214017743586>
- Baylor, R., Fatehi Y. K., & Esper, H. (2020). *Advancing the use of developmental evaluation: A summary of key questions answered during a multiyear study of developmental evaluations implemented at USAID*. DEPA-MERL Consortium. USAID. <https://www.usaid.gov/GlobalDevLab/MERLIN/DEPA-MERL/analysis-across-pilots>
- Baylor, R., Fatehi, Y. K., & Esper, H. (2019). Key lessons from an attempted developmental evaluation pilot. DEPA-MERL Consortium. USAID. <https://www.usaid.gov/GlobalDevLab/MERLIN/DEPA-MERL/bureau-food-security/bfs-learning-report>
- Beer, T. (2019). *Are you really ready for developmental evaluation? You may have to get out of your own way*. Center for Evaluation Innovation. <https://www.evaluationinnovation.org/in-sight/are-you-really-ready-for-developmental-evaluation-you-may-have-to-get-out-of-your-own-way/>
- Beer, T., Lee, K., & Ward, K. (2019). *Developmental evaluation: Rewards, perils, and anxieties* [Webinar]. Gender and Evaluation. <https://gendereval.ning.com/events/webinar-developmental-evaluation-rewards-perils-and-anxieties>
- Blanchet-Cohen, N., & Langlois, M. (2010). *DE 201: A practitioner's guide to developmental evaluation*. International Institute for Child Rights and Development, University of Victoria.
- DEPA-MERL Consortium. (2019a). *Implementing developmental evaluation: A practical guide for evaluators and administrators*. U.S. Agency for International Development. https://www.usaid.gov/sites/default/files/documents/15396/ImplementingDE_Admin_20.pdf
- DEPA-MERL Consortium. (2019b). *Implementing developmental evaluation: A practical guide for funders*. U.S. Agency for International Development. https://www.usaid.gov/sites/default/files/documents/15396/ImplementingDE_Funders_20.pdf
- Fatehi, Y. K., Baylor, R., Esper H. (2018). *A study of the family care first in Cambodia developmental evaluation*. DEPA-MERL Consortium. USAID. <https://www.usaid.gov/GlobalDevLab/MERLIN/DEPA-MERL/family-care-first/final-version>
- Gamble, J.A.A. (2008). *A developmental evaluation primer*. J.W. McConnell Family Foundation. <https://www.betterevaluation.org/sites/default/files/A%20Developmental%20Evaluation%20Primer%20-%20EN.pdf>
- Gold, J., Wilson-Grau, R. W., Fisher, S., & Otoo, S. (2014). *Cases in outcome harvesting: Ten pilot experiences identify new learning from multi-stakeholder projects to improve results*. The World Bank.

- Hayes, H., Witkowski, S., & Smith, L. (2016). Failing forward quickly as a developmental evaluator: Lessons from year one of the LiveWell Kershaw journey. *Journal of MultiDisciplinary Evaluation*, 12(27), 112–118.
http://journals.sfu.ca/jmde/index.php/jmde_1/article/view/435
- Lam, C. Y. (2016). *A case study on a design-informed DE*. QSpace.
<http://hdl.handle.net/1974/14929>
- Lam, C. Y., & Shulha, L. M. (2015). Insights on using developmental evaluation for innovating: A case study on the co-creation of an innovative program. *American Journal of Evaluation*, 36(3), 358–374.
<https://journals.sagepub.com/doi/abs/10.1177/1098214014542100>
- Langlois, M., Blanchet-Cohen, N., & Beer, T. (2012). The art of the nudge: Five practices for developmental evaluators. *The Canadian Journal of Program Evaluation*, 27(2), 39.
https://www.researchgate.net/publication/289246043_The_art_of_the_nudge_Five_practices_for_developmental_evaluators
- Miller, R. L. (2010). Developing standards for empirical examinations of evaluation theory. *American Journal of Evaluation*, 31(3), 390–399.
<https://journals.sagepub.com/doi/10.1177/1098214010371819>
- Mintzberg, H. (2007). *Tracking strategies: Toward a general theory*. Oxford Press.
- Patton, M. Q. (1994). Developmental evaluation. *American Journal of Evaluation*, 15, 311–319.
<https://journals.sagepub.com/doi/10.1177/109821409401500312>
- Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. Guilford Press.
- 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.
<https://journals.sagepub.com/doi/10.1177/1098214015626295>
- Patton, M. Q., McKegg, K., & Wehipeihana, N. (Eds.). (2015). *Developmental evaluation exemplars: Principles in practice*. Guilford Press.
- Patton, M.Q., & DEPA-MERL Clinic #9. (2019). *How developmental evaluation informs learning and adaptation* [Webinar]. Not available on a public domain.
- Stufflebeam, D. L. (2001) The metaevaluation imperative. *American Journal of Evaluation*, 22(2):183-209.
<https://journals.sagepub.com/doi/10.1177/109821400102200204>
- Wilson-Grau, R. (2015). *Outcome harvesting. Better Evaluation*.
https://www.betterevaluation.org/en/planning/approach/outcome_harvesting