Article

American Journal of Evaluation
2023, Vol. 44(4) 629-648

DOI: 10.1177/10982140211050912
journals.sagepub.com/home/aje
Sage

Research on Evaluation Influence in India: Theorizing Beyond Process and Results to Design

Melissa Rae Goodnight[1]

Abstract

This study analyzes evaluation influence theories to understand their unified contributions to a conceptual framework for research on evaluation influence in non-western contexts. Specifically, these theories are analyzed according to their usefulness for interpreting the consequences of the Annual Status of Education Report (ASER)—a cyclical, large-scale educational monitoring and evaluation effort in India. An ethnographic case study of ASER facilitates an empirical investigation of the field's evaluation influence theories and uncovers a new source of influence—evaluation design. The study's findings challenge a common belief in the field: that evaluation process and results are the two fountainheads of evaluation impact. Design, as a distinct third source of influence, gives rise to a novel form of evaluation consequence—design diffusion. The phenomenon of design diffusion illuminates why ASER's model for evaluation has been adopted by several other Global South nations.

Keywords

India, evaluation theory, evaluation design, evaluation influence, research on evaluation

The purpose of this article is to explore *evaluation influence* as a conceptual framework for conducting *research on evaluation* (RoE) in non-western contexts. The explanatory power of evaluation influence is examined via its usefulness for interpreting the consequences of the *Annual Status of Education Report* (ASER). ASER is a cyclical, large-scale educational monitoring and evaluation (M&E) effort in India. An ethnographic case study of ASER facilitates this empirical investigation of the field's evaluation influence theories. These theories have not been analyzed collectively for their unified contributions to a conceptual framework for research on evaluation influence (RoEI). Through employing the RoEI framework, the ASER case study uncovers a new source of evaluation influence beyond an evaluation's process and results—its evaluation design. The field has long

[1] PhD in Education, Assistant Professor in Educational Psychology and Global Studies, University of Illinois at Urbana-Champaign

Corresponding Author:
Melissa Goodnight, 1310 S. Sixth St., Education Building 210F, Champaign, IL 61820, USA.
Email: mrg@illinois.edu

regarded evaluation process and results as the two fountainheads of evaluation impact. Design, as a distinct third source of influence, gives rise to a novel form of evaluation consequence: *design diffusion*. The phenomenon of design diffusion explains a key consequence of ASER: why its model for the citizen-engaged evaluation of education systems has been adopted by numerous other Global South countries.

The article proceeds with a brief background on research on evaluation, and monitoring and evaluation in Global South contexts. Then, evaluation influence theories are analyzed, and their dimensions synthesized into one conceptual framework for guiding RoE. Next, the ASER ethnographic case is presented focusing on two dimensions of evaluation influence from the RoEI framework: source and intention. Drawing on ASER case data, a new source of evaluation influence (evaluation design) and a new form of influence (design diffusion) are explained in a section on ASER's design. The article concludes with lessons learned from applying the RoEI framework and suggestions for future research on evaluation influence.

Research on Evaluation to Improve Theory

A growing focus within the field of evaluation is conducting empirical studies of evaluations to better understand their features, implementation processes, direct impacts, and broader influences (Coryn et al., 2016). This constellation of research foci and their corresponding inquiry activities are generally referred to as research on evaluation or RoE. Henry and Mark (2003b) asserted the benefit of evaluation studies, including *research on evaluation outcomes* as one of six types of RoE*,* for developing the field's cumulative knowledge. More recently, Alkin and Christie (2019) explained that the areas needing greatest attention in the empirical literature on evaluation were (1) evaluation theory and (2) *theory into practice* in which researchers use data to understand how evaluation theories are translated into actual evaluation practices.

In terms of methodologies well suited for RoE, case studies of evaluations can illuminate problems and potentialities with applying particular theories or approaches in concrete contexts with real stakeholders. Case studies help evaluators reflect on existent evaluation theories—not only the mass of *prescriptive* theories (i.e., models) outlining how evaluations should be done, but also the comparatively fewer *descriptive* theories developed to help evaluators better explain the nature of evaluation and understand its consequences (Alkin, 2013; Christie & Alkin, 2013). Evaluation influence theories are descriptive of an evaluation's effects and, despite the rarity of doing so (especially in non-Western contexts), influence theories can be explored through empiricism (e.g., Gildemyn, 2014). Evaluation case studies provide the means to critique whether a theory is explanatory of the observed consequences of an evaluation. Thus, the ASER case study facilitates better understanding of evaluation influence theories' usefulness for interpreting M&E outcomes in India.

Research on Monitoring and Evaluation in the Global South

Nearly 40 years ago, Nevo (1982) made the case for systematically studying evaluations in international contexts through empirical research out of concern about the ill-fitting application of US-specific evaluation theories and models to other countries. Research on evaluation outside of Western or Global North contexts is comparatively underrepresented in the expanding RoE literature even though it is widely understood that cultural context is a fundamental aspect of any evaluation endeavor (Chouinard & Cousins, 2009). Substantive literature on the importance of cultural responsiveness in conducting valid evaluations proves that good theories, methods, and practices are not universally appropriate but contextually and population specific (e.g., Hood, 2004; Hopson, 2009; Kirkhart, 1995, 2010). Global North worldviews and approaches are often uncritically employed in designing and conducting evaluation for "developing" countries with consistently harmful

consequences for programs, policies, and stakeholders (Ofir & Kumar, 2013). In fact, Chilisa, Major, Gaotlhobogwe, and Mokgolodi (2016) contend that "evaluation has become the worst instrument of epistemological imperialism" in the Global South by determining what "should be considered real program outcomes, what knowledge measures that reality, and what values support the evaluation practice" (p. 314). These marginalizing processes are wrongly assumed to be objective and culturally universal. For the field of evaluation to best serve the public good in a diversity of countries, a more expansive, nuanced and culturally insightful view of evaluation context and theory is essential. Today, monitoring and evaluation is widely viewed as a solution for ensuring the effective delivery of social programming and the transparent operation of national public systems like education and public health (Gildemyn, 2014). M&E efforts need to be grounded in sound evaluation theories, methods, and practices that are responsive to the various national and cultural contexts in which they take place. A crucial approach for developing and promoting better theories and practices is conducting more research on evaluation in Global South countries to (1) study the components and practices of M&E efforts that originate locally; (2) test, refine, and expand current evaluation theory frameworks for appropriate use in these evaluation contexts; and (3) responsively develop evaluation theory from the ground up.

Theories of Evaluation Influence

> *The term influence (the capacity or power of persons or things to produce effects on others by intangible or indirect means) is broader than use, creating a framework with which to examine effects that are multidirectional, incremental, unintentional, and noninstrumental, alongside those that are unidirectional, episodic, intended, and instrumental (which are well represented by the term use).* (Kirkhart, 2000, p. 7)

Theories of evaluation influence encourage researchers of evaluation to cultivate a holistic view in investigating evaluation's implications for stakeholders and the *evaluand* (i.e., program, policy, or public system under evaluation). Kirkhart's (2000) "integrated theory of influence" builds upon the field's narrower notion of *evaluation use* to conceptualize a unified, fuller understanding of evaluation's meaning. Henry and Mark (2003a) construe evaluation influence as a way to understand evaluation's consequences by theorizing the change mechanisms by which an evaluation has influence. Hall (2004) urges a step back from thinking about the effects of an evaluation's formal conduct and results to highlight the anticipation of an evaluation as a major influence on social programs and stakeholder behavior.

The concept of influence is not without its critics (e.g. Alkin & King, 2017; Alkin & Taut, 2003; King & Alkin, 2019; Nunneley, King, Johnson, & Pejsa, 2015). Nevertheless, this article asserts that a synthesized framework of evaluation influence theories importantly builds upon the legacy of evaluation use theorists' insightful research on the nature of evaluation and its consequences (Alkin & King, 2016, 2017), and it also allows for the consideration of a broader spectrum of effects than those typically associated with the conventional aims of evaluation and more generally described as "use". This expanded perspective is key in discovering novel ways in which individual evaluations may have lasting impacts on social programs and social inquiry.

Three Dimensions for Conceptualizing Evaluation Influence

Kirkhart's (2000) theory highlights the relationship between three intersecting dimensions of evaluation influence: *source*, *intention*, and *time*. The first dimension, the source of influence, describes whether an evaluation's influence emanates from its *process*, i.e., "the process of conducting the evaluation itself," or its *results*, i.e. the information produced during an evaluation—its data, findings, and recommendations (pp. 8–10). Thus, two kinds of evaluation influence, *process-based influence* and

results-based influence, are broadly conceived (Kirkhart, 2011). The second dimension describes "the extent to which evaluation influence is purposefully directed, consciously recognized, and planfully anticipated" (Kirkhart, 2000, p. 11); in other words, the dimension examines whether the effects of an evaluation's process or results are *intended* or *unintended* according to the evaluation's purpose and goals that are embedded in the evaluation's design. Three aspects of intention are identified: (1) the type of influence desired or anticipated, (2) who is to be influenced by the evaluation, and (3) the processes, findings, or people that/who are expected to exert influence. The third dimension of time pertains to the "chronological or developmental periods in which evaluation influence emerges, exists, or continues" (Kirkhart, 2000, p. 9). The sub-dimensions of time—*immediate*, *end-of-cycle*, and *long-term*—identify "the need to recognize influence during and immediately following the evaluation cycle as well as effects that are visible in the future" (2000, p. 9). In sum, Kirkhart's theory offers a framework for theorizing an evaluation's immediate, end-of-cycle, and long-term consequences as intended or unintended manifestations of the evaluation's process or results. Kirkhart cautions that her influence theory's sub-dimensions or categories should probably be treated as continua because the realities of an evaluation's influence are in the "gray areas that fall between" (2000, p. 8). While providing a concrete set of intersecting analytical categories, Kirkhart has increasingly warned against the dimensions' use toward dichotomous thinking and rigid interpretations (e.g., 2011). Moreover, it is plausible that an evaluation may have relatively little influence or contrastingly, multiple influences—some intended and some not—emanating from different sources and occurring at various times.

Social Betterment and Levels of Influence

> *Theories of evaluation influence cannot and should not attempt to direct attention to all of the possible consequences of evaluation. A theory of evaluation influence should include a way of directing attention to some possible effects (e.g., actions affecting service delivery) but not others (e.g., evaluators receiving paychecks). We propose that a theory of evaluation influence should focus on the* ***subset of evaluation consequences that could plausibly lead toward or away from social betterment*** *(*Henry & Mark, 2003a*, p. 295)*

Henry and Mark (2003a) and Mark and Henry (2004) adopt Kirkhart's terminology of evaluation influence but suggest that influence should be understood via identification of mechanisms of change (i.e., mediating factors). They state that comprehending influence is essentially synonymous with discovering "the change processes through which evaluation affects attitudes, beliefs, and actions" and "the interim outcomes that lie between the evaluation and its ultimate goal—social betterment" (Henry & Mark, 2003a, p. 293). Thus, evaluation influence is comprised of dynamic processes that inspire consequential, interim changes toward a long-term objective of social betterment—a term commonly used and loosely defined throughout evaluation literature that is problematic because it raises questions about *whose* values define better (King, 2016). Henry and Mark (2003a) define social betterment as a state in social conditions "better than the state that existed before, as judged through deliberation and by public opinion" within a democratic society (p. 295). An evaluation's interim outcomes, which are actually change processes, characterize its particular influence pathway and occur at three levels of influence: the individual, interpersonal, and collective (Henry & Mark, 2003a). Individual-level influences are characterized as instances when the evaluation motivates changes predominantly within a person. Interpersonal-level influences refer to changes that are relational, meaning chiefly to processes or outcomes taking place between people. Collective-level influences are changes to processes or outcomes that are institutionalized, happening within public or private organizations. The actual change processes of an evaluation occurring at different levels constitute a specific evaluation's local theory of influence. What Henry and Mark

(2003a) describe is offering a general menu of change mechanisms that are likely to occur across various contexts and together constitute a general theory of evaluation influence from which local theories can be developed. They modify Kirkhart's chiefly descriptive theory of influence—that is an open framework for analyzing any number of evaluation consequences—by infusing evaluation influence with the morality of social betterment, and in doing so, they propose concentrating on substantive social changes at the individual, interpersonal, and collective levels. Henry and Mark's theory of influence is less purely descriptive; they argue certain criteria should guide where attention be directed in analyzing evaluation's effects. Henry and Mark present influence as both evaluation's end result and interim changes toward the grand impact of social betterment.

Anticipating Evaluation

> *Evaluation as a concept is well entrenched in the non-evaluation world and as a result influences behavior beyond the scope of formal evaluation practice. To fully understand influence, our lens of inquiry must be broadened to understand the myriad of conceptual understandings, emotional reactions, and individual appreciations resident in those impacted by evaluation and present in their thinking long before we as evaluators arrive on the scene (*Hall, 2004, paragraph 3).

Hall (2004) has proposed the existence of *anticipatory influence* in evaluation to address the sociocultural and political contexts of evaluations. Accordingly, stakeholders' predispositions on evaluation mold how a future evaluation will unfold, be experienced, and be received. These stakeholder feelings about its fairness or usefulness can pre-date any formal commission and conduct of evaluation. Sometimes, the anticipation of evaluation can significantly shape social programs irrespective of the methodologies eventually employed to conduct the evaluation. In other cases, when methods are clearly "pre-specified," Hall (2004) notes how the "anticipatory impact on behavior can even become prescriptive" in how a social program is envisaged and implemented (paragraph 5). Hall's theorization of anticipatory influence also supplements the field's understanding of how time operates in terms of evaluation influence by expanding the time dimension first proposed by Kirkhart to sub-dimensions of anticipatory, immediate, end-of-cycle, and long-term.

Finding the Language to Tell an Evaluation Story

Language in evaluation, as in other fields, is important because it shapes people's thinking and creates a foundation for the exchange of ideas. The field's language regarding evaluation use and influence fundamentally structures how evaluators are trained to envision and plan for an evaluation's impact. This envisioning and planning are fundamental to the intended aspects of an evaluation conceptualized in its design. The language of use and influence also directly impacts how evaluation researchers study the consequences of evaluation.

Given the evaluation field's organic interdisciplinary development, there are unsurprising conceptual challenges with its terminology that evaluation influence theories aim to address. Kirkhart's (2000) proposal of an integrated theory of influence is an attempt to overcome the fragmentation of past conceptualizations of evaluation's effects that have been too narrowly and simplistically characterized by the priorities of instrumental, results-oriented use and have constrained appraisals of evaluation's influence and full meaning. Regarding the limitations of evaluation use, Henry and Mark (2003a) assert that "simply adding to categories and definitions that appear under the umbrella of use is no longer sufficient for the field to understand and thoroughly examine the consequences of evaluation" (p. 309). In particular, the descriptors employed in defining different types of evaluation use indicate unreliable language conventions that have marred the clarity of the field's theorizing: *process use*, *symbolic use*, *imposed use, instrumental use*, and *conceptual use* are all terms meant

to connote some type of evaluation effect or utilization though they do not fit together within one typology. In fact, these multiple "taxonomies of use" describe disparate things (Mark, 2011, p. 111): (1) the *source* of an evaluation's effects (e.g., the evaluation process in process use); (2) the *motivation* behind an evaluation's conduct and utilization (e.g., the desire to legitimize decisions in symbolic use); and (3) the *form* of an evaluation's utilization (e.g., direction action in instrumental use) and the *form* of an evaluation's effect (e.g., enlightenment in conceptual use). These types of evaluation use are not mutually exclusive from one another (Mark, 2011), reflecting discordant intellectual scaffolds for describing use that if harmonized could contribute to a complex understanding of evaluation's multiple consequences. The dimensions (source, intention, time) of Kirkhart's theory do not address the *forms* that evaluation influence takes nor the *motivations* that lead to engaging in evaluation. Kirkhart's intention dimension infers motives are present in an evaluation but focuses on whether an evaluation's consequence was planned rather than identifying the reasons for evaluation action.

Subsequent conceptualizations of influence have addressed additional evaluation attributes. Henry and Mark (2003a) discuss motivation in their articulation of social betterment, which proposes an ideal motivation for the change processes that evaluations can inspire and that can lead to influence. Hall's description of anticipatory influence focuses on the psychology of evaluation and provides a unique meditation on motivation in its consideration of the mental and emotional dimensions of evaluation for various stakeholders. Hall does not represent any stakeholder feelings or reactions as legitimate or illegitimate as is the case with symbolic and imposed use, where the motivations of stakeholders are construed as problematic, manifesting into types of evaluation *misuse*. Henry and Mark (2003a) outline several specific forms of evaluation influence (e.g., *priming* and *agenda setting*) occurring at different levels (i.e., individual, interpersonal, and collective). Diffusion is a form of collective-level influence that entails an "adoption of the policy, program or practice being evaluated" in an environment outside of where the evaluation took place (p. 305). Diffusion indicates evaluations can have broad impact beyond their intended stakeholders and contexts—this concept will be returned to later in the article as diffusion becomes a vital foundation for understanding ASER's influence.

More recently, Mark (2011) portrays evaluation use as terminology that presents dismembered chapters of an evaluation's *story*: process use is somewhere in the beginning, imposed use illuminates the middle or end, and instrumental use and conceptual use reflects the end. Henry and Mark's theorizing of influence pathways is an attempt to overcome patchy accounts of an evaluation by tracing it from conceptualization to long-term consequences through a causal chain of influences that offers a coherent narrative of an evaluation's impact. Stepping back to survey the influence-use relationship, studying evaluation influence does not diminish the contributions of the use literature toward the field's deeper understanding of evaluation's nature and consequences; rather, evaluation influence broadens the scope of what researchers of evaluation contemplate in telling an evaluation's story.

Evaluation Context: History, Language, Politics and Culture

A storytelling metaphor is valuable for research on evaluation influence in how it draws attention to the issue of the setting for an evaluation's action—its context. Hall's (2004) anticipatory influence emphasizes that the story of an evaluation (and its consequences) begins long before an evaluation has been planned. Hall spotlights the importance of a context's *history* in how evaluation will ultimately unfold. Understanding this history and other features of the evaluation context—language, politics, and culture—determine an evaluator's competence and this competence essentially shapes the evaluation and its influence.

An evaluators' knowledge of *language* within the evaluation setting and across various stakeholder groups matters; good communication eventually determines whether an evaluation is

persuasive enough to have influence (Hall, 2004). Hall (2004) asserts that using language purposefully is essential to the persuasiveness of an evaluation and persuasion is a necessary prerequisite to influence. Culture certainly shapes the use, interpretation, and persuasiveness of language, so communicating "with fidelity across group boundaries" is a vital consideration for evaluations, particularly in multicultural settings (Hall, 2004, paragraph 10). Henry and Mark (2003a) also identify persuasion—an individual using direct communication in an attempt to change another person's view—as a change mechanism of influence at the interpersonal level.

Examining culture is not optional in telling a truthful story about evaluation or its influence. Hall (2004) explains that the field's influence theories are weakened by not recognizing the significance of culture in evaluation: "ignoring cultural context won't reduce its impact, ignoring cultural context will however, result in impact being either unexplained or misinterpreted" (Hall, 2004, paragraph 15–16). Kirkhart (2011) also emphasizes that "evaluation influence must be understood and studied as a cultural phenomenon" that manifests both individually and organizationally across multiple settings comprising an evaluation (p. 73). Culture is a fluid, complex, and untidy concept—it defies easy operationalization in research on evaluation and requires that evaluators and RoE scholars approach its conceptualization, navigation, and analysis with *humility* (Kirkhart, 2011). The treatment of culture in evaluation is central to issues of social justice: evaluators who are inattentive to culture frequently possess worldviews uninterrogated for their biases and preferences and consequently, uncritically reflect majoritarian (e.g., "white, heterosexual, middle class, able bodied" and Global North) "norms, values and assumptions" in the theories and methods they develop and employ in evaluations (Kirkhart, 2011, p. 74 & 76). Likewise, cultural views mold evaluators' understanding of and planning for influence. Asserting there are no culture-free interpretations of evaluation, Kirkhart (2011) offers four observations on what culture teaches the field about influence: (1) "avoid dichotomous thinking" and rigid categorizations that simplify lived complexities; (2) reconsider dominant and exploitative practices around "ownership", "knowledge", and community autonomy; (3) seek understanding of the connections between an evaluation's consequences, equity, and social justice; and (4) acknowledge the "importance of history" in evaluation contexts including communities' beliefs and experiences related to programs, evaluations, and interpretations of their significance (p. 81–82). These observations offer insight on values that may guide interpretations of evaluation influence motivated by a social justice vision.

Additionally, the politics of an evaluation setting markedly impacts the unfolding of its unique story. Analyzing the role of political factors unifies classic research on evaluation use with recent literature on evaluation influence. Carol Weiss consistently highlighted politics as a pervasive factor in evaluation use, locating its significance in the formal policy and institutional infrastructures that shape evaluation contexts and in the diverse motivations, beliefs, and assumptions of individuals with varying degrees of decision-making power within these infrastructures (e.g., Project Team, 2006; Weiss, 1999). The multilevel lens offered by Henry and Mark supplements this appraisal of politics in interpreting an evaluation's influence even as their proposed levels seem artificially bounded: certainly, political influences within evaluation contexts can operate vertically where local, regional, national, and global dynamics mold one another; it would seem the same can be true of politics across individual, interpersonal and collective levels of influence. Hall (2004) expresses his initial interest in evaluation was rooted in a commitment to defending social programs, which benefited African American communities, against the racist politics that have historically shaped the values guiding their evaluation. Hall infers these different values guided his social justice vision for interpreting evaluation's influence that emanated from a culturally responsive lens. His vision of evaluation influence is attentive to factors of race, ethnicity, language, and culture as well as power (i.e., politics).

A Conceptual Framework for Research on Evaluation Influence

For those engaged in evaluation research, synthesizing the work of evaluation influence scholars provides a rich RoEI conceptual framework for interpreting the often prolonged and complicated story of

an evaluation and its impact. Their theories reveal seven dimensions for telling an evaluation story (Table 1)—source, intention, time, motivation, level, form, and context—which inform this article's study of ASER's influence.

Ethnographic Case Study of the ASER Evaluation

A ten-month multi-state ethnographic case study of the 2014 cycle of ASER was conducted in India with field research in villages, schools, partner organizations, and the evaluation organization's (ASER Centre) central offices. The study employed (1) participant observations, (2) semi-structured interviews, and (3) document analysis to examine ASER's purpose and goals, design choices, implementation, dissemination of findings, and influence (Goodnight, 2017a).[1] In addition to ASER staff and participants, members of India's "policy-shaping community" (Cronbach, 1982) were interviewed. The next two sections provide an analysis of ASER guided by the dimensions of source and intention with support from other dimensions (e.g., time, level, and motivation) outlined in Table 1. The resulting analysis reveals evaluation design as an overlooked source of influence in an account of ASER's influence. The third section explores the dimension of context (e.g., politics, history, language, and culture) in relation to ASER's design.

Source

ASER Process

The Delhi-based nongovernmental organization ASER Centre conducts ASER in all of India's rural districts with the help of approximately 500 partner organizations and 25,000 volunteers countrywide to investigate the conditions of government primary schools and record the enrollment status and learning levels of school-age children. The execution of ASER encompasses a large range of activities—training sessions, travel to villages, data collection, monitoring, data recheck and aggregation, data analysis, and report writing and printing. Table 2 depicts the phases of the ASER process, which involves coordinating hundreds of ASER Centre's own staff as well as volunteers and partners. Executed through a hierarchical participatory structure, ASER is directed from the top by ASER Centre national team members who oversee state teams and orchestrate logistics between states and the national office. Meanwhile, state team members liaison with partner organizations and supervise master trainers and volunteers executing ASER in rural districts within their respective states.[2] With the help of partner organizations, master trainers coordinate logistics for ASER at the district-level: training workshops, monitoring volunteers' data collection, and conducting data quality rechecks. Two volunteers travel to each sampled village to map and survey the village,

Table 1. Evaluation Influence Dimensions.

	Dimension	Sub-Dimensions or Categories
1	Source	Process, Results
2	Intention	Intended, Unintended
3	Time	Anticipatory, Immediate, End-of-Cycle, Long-term
4	Motivation[a]	Social Betterment, Social Justice for Culturally Marginalized Groups, Misuse
5	Level	Individual, Interpersonal, Collective
6	Form	e.g., Elaboration, Priming, Justification, Persuasion, Standard Setting, Diffusion[b]
7	Context	Language, History, Culture, Politics

[a]Negative motivations are discussed in evaluation use literature as misuse and indicated by Hall (2004) as harm.
[b]Henry and Mark (2003a) and Mark and Henry (2004).

Table 2. The Phases of the ASER Process.

Phase	Description of Activities	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
I	Forming partnerships and recruiting master trainers and volunteers										
II	Training sessions (national, state and district) on conducting ASER										
III	Collecting data and monitoring in the field (villages, schools and houses)										
IV	Rechecking data (desk, phone and field verifications) and resurvey										
V	Aggregating and analyzing data and report writing										
VI	Publishing and distributing report and hosting national media release										
VII	Hosting state-level releases and meeting with partners										

The timeline depicted in this table reflects the phases of the ASER process during ASER 2014. This table is adapted from a table in ASER Centre's (2014) "Quality Control Framework" (p. 6).

survey the government primary school, and select and survey 20 households (see ASER, 2015, pp. 313–315 for sampling design information). During household surveys and testing, volunteers inquire about the family's economic status, number of school-age children, and parents' educational attainment and then administer reading, arithmetic, and English language tests to all children (5–16 years old). This survey and testing process is repeated for 30 villages in each district. A tight 100-day timetable from "field to report" is a key feature of ASER's process intended to produce up-to-date results that can inform decision-making in real time.

ASER Results

ASER's final product is a thick, widely publicized report that provides current findings on rural primary education, highlighting issues related to India's universalization efforts and its educational governance (ASER Centre, 2015). ASER data tables illustrate by state and district, and disaggregated by gender, grade level, and enrollment status, children's ability to read simple texts, do basic math, and demonstrate beginner English skills. Currently, ASER is conducted every two years and offers the only biennial publicly available figures on learning pan-India; from 2005–2014, ASER was annual with data publicly released every January.

Intention

The intention dimension prompts a consideration of intended versus unintended consequences from an evaluation, and these two directions for inquiry organize the following discussion of the ASER case. Embedded in this discussion of intention are the dimensions of time, motivation, form, and level from Table 1.

Mapping Intended Influence Pathways

Regarding intended consequences, ASER's intended influence pathways can be mapped to depict its *proposed* local theory of evaluation influence. Mark (2011) recommended such an exercise as important to explorations of evaluation influence and similar to depicting a program theory through logic models.[3] In depicting ASER's intended influence pathways, two kinds of evaluation influence appear possible guided by the source dimension: (1) *process-based influence*—relating to the effects of the ASER process on institutions and individuals who participate in ASER, and (2) *results-based influence*—relating to the effects of ASER results on policies, practices, and public dialogue. Through empirical analysis of ASER via observations, protocol analysis, and ASER architect interviews, four intended influence pathways were identified with two main goals—grassroots action and policy change—emanating from these different sources. Figure 1 illustrates ASER's intended influence pathways.

Process-based intended influence pathways. The first goal of ASER—grassroots action—is intended to stem from participation in its process. Figure 1 shows two intended pathways (As denoted by the superscripts a and b) of ASER's process-based influence—one related to a trajectory of individual grassroots action and the other to coalitional (i.e., group) grassroots action. Coalitional action seemed to rely on an interim change at the individual level (i.e., increased individual awareness) that leads to one at the group level (i.e., enhanced group awareness). ASER's process is meant to generate individual learning and awareness as well as organizational or community awareness. Community-level action undertaken by individuals or groups reflects ASER's process-based influence, which may be intermediately evaluated through investigating in what ways ASER's process immediately stimulated volunteers', partners', or villagers' awareness about government schools, children's learning, or education more broadly. The end-of-cycle objective in these intended influence pathways is local deliberation and organizing around education (i.e., individual or group deliberation), and the long-term objective is ongoing, well-orchestrated action serving an ongoing commitment to the cause of education. The ultimate goal of ASER's process (i.e., its motivation) can be viewed as social betterment—leaving an area's education in a better state as judged by local citizens—via improved, sustained educational practices.[4]

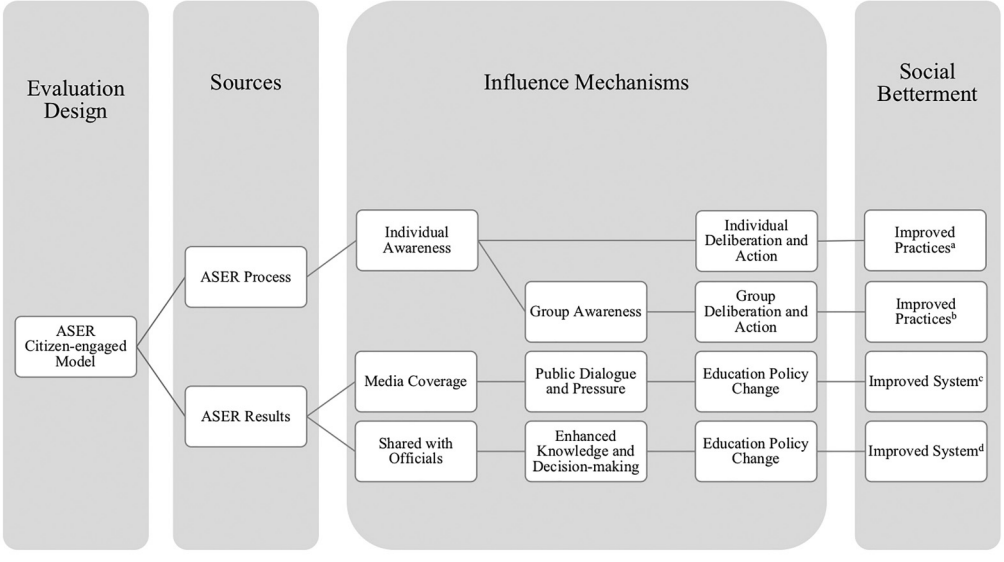

Figure 1. ASER's model of intended influence pathways.

Results-based intended influence pathways. Two intended influence pathways are rooted in ASER's goal of policy change and reflect its results-based influence (pathways c and d). In pathway c, the means by which ASER results are intended to affect policy change are via prompting public dialogue and pressure on government officials, so ASER's influence on policy may be intermediately evaluated through investigating to what extent ASER's results prompt public deliberation about learning and the primary education system, and put pressure on government toward making policy changes. Consequently, an end-of-cycle objective for increasing ASER's results-based influence is widespread media coverage of ASER results with a long-term objective of government making education policy changes (e.g., budgetary reforms).[5] The ultimate goal of ASER's results can be perceived as social betterment: an improved education system as a result of constructive policy change. Pathway d in Figure 1 illustrates various interim objectives of officials' enhanced knowledge and decision-making with the same ultimate goal of an improved education system. This last pathway also begins at the end of ASER's cycle with directly sharing ASER results with key officials but seems to reflect a different assumption about officials' motives—that they will be amenable to reviewing ASER results, to assimilating them into their knowledge of educational issues, and subsequently, to considering them in their decision-making on educational policies.

Intended pathways versus actual influence. ASER's intended influences—model pathways for achieving ASER's specific vision of social betterment—are central to its design. Mapping ASER's intended influence pathways creates a visual articulation of its underlying theory that can then be used for interpreting its manifest influence pathways. Its manifest pathways can be observed through RoE data to surmise how ASER has achieved (or not) what it envisioned.[6] Like a logic model depicting a program theory, mapping intended influence pathways gives researchers of evaluation an initial understanding of how evaluators—or the architects of M&E —foresee it having influence.

A model of intended influence pathways can provide some methodological direction to RoE researchers by indicating areas of the evaluation where data capture is most important for investigating achievement of intended impact. Data collection strategies informed by intended influence pathways may yield data that can better illuminate the mechanisms of change. For instance, in ASER's intended pathway[a], collecting qualitative data from ASER volunteers could help the RoE researcher determine whether increased individual awareness (via participating in ASER's process) takes the form of *salience* (the issue of learning is more important), *elaboration* (the issue of learning is better understood), *attitude change* (new attitudes toward the issue of learning are formed), or some combination of these, or none of these (Henry & Mark, 2003a). Thus, proceeding from intended pathways, researchers may explore use and influence through empirical evidence, but these intended models of influence do not take away the complexity of analyzing actual effects. For instance, ASER data has been cited for several years in Government of India's Ministry of Finance's Economic Survey reports (see Government of India, 2014, as an example). It is important to consider the meaning of this use by studying the extent to which it reflects something more in accordance with ASER's desired changes—such as prior, ongoing, or future use in the government's decision-making (i.e., *policy change*) (Henry & Mark, 2003a). In other words, does publicly citing ASER results in the report reflect ASER's broader influence on knowledge and decision-making in the Ministry? Does enhanced decision-making lead to positive policy changes for the education system? ASER's larger meaning cannot be understood by disentangling the observed episodic uses of its results or processes from a fluid, comprehensive conceptualization of its influence.

Intended pathways as ideal models for results-based influence or process-based influence are oversimplified and excessively linear in comparison to the messiness of how influence actually unfolds. In reality, it can be difficult to decipher which source of the evaluation has prompted the effects that lead to evaluation influence (Gildemyn, 2014). This analytical fuzziness in evaluating the veracity of

intended influence pathways points to challenges in constructing real influence pathways from the analysis of data, especially for an evaluation at ASER's scale. Ultimately, determining whether ASER's intended pathways of influence translate to actual influence is a complex question requiring extensive empirical study to determine how and why ASER matters to the Indian education system.

There are issues with being strictly focused on testing ASER's underlying theory by looking at its intended influence pathways. One issue is how the construction and testing of ASER's intended influence pathways may operationalize evaluation influence in a way that diminishes its power as a *sensitizing concept* by reducing its abstraction (Patton, 2020), and thus, its accuracy for detecting impact across the varying cultural contexts of ASER in India. Another issue is the oversight of unintended consequences (and potential minimization of negative ones). At best, the intended influence pathways offer a method of reflection that aids the RoE researcher in looking beyond an evaluation's anticipated influence to identify unintended influence emerging from data.

Investigating Unintended Influence in ASER

Unintended consequences of an evaluation may be negative, positive, both, or ambiguous in their contribution to the change (e.g., social betterment) envisioned. Methodological issues arise in studying unintended influence including (1) knowing when and where to look for unintended effects, and (2) foreseeing what data collection strategies might best capture these effects, so they are perceptible to the RoE researcher. In interviews with ASER's architects, a commonly raised challenge was knowing in what timeframe they should follow-up with volunteers, partner organizations, villages, or government officials to learn about ASER's impact. How to determine what is a reasonable amount of time to let pass before trying to investigate influence? While studying ASER's long-term influence is significant to ascertaining its overall meaning, serious challenges are posed by its scale and the issue of time. Field observations and interview data reveal ASER is an intense process for its participants, so revisiting the possibility of extended impact at a future date seems logical as participants may take months or years to assimilate and employ all that they have learned. Though tricky to design, such an investigation also benefits the evaluation field's literature because, while recognized as crucial, there are few empirical studies of long-term influence (Kirkhart, 2000). Thus, in taking a protracted and expanded view of ASER's influence, this study's analysis of interviews, fieldnotes, and documents revealed an unintended type of influence (not captured in ASER's intended influence pathways). The next section details this long-term unintended influence: how ASER's evaluation design became a source of influence within a global policy-shaping community eager to find solutions for measuring learning at a large scale.

Evaluation Design—A New Source of Influence

The American Evaluation Association defined an evaluation's *design* as the integration of "evaluation theories, approaches, and methods to achieve a set of intended purposes in a specific context."[7] The next three sections describe ASER's design as a citizen-engaged model for evaluating education systems that heavily reflects India's evaluation context (e.g., its politics) but has gained global prominence and been translated to other countries.

ASER's Design

Theory. The citizen-engaged aspect of ASER is born not just out of a logistical need to overcome data collection challenges but also a philosophical stance on popular engagement in education. ASER directly advocates public ownership of the social systems that affect people's everyday lives. ASER's architects intended to realize M&E's *democratic* potential (House & Howe, 2000) by

encouraging information sharing and deliberation to advance the public good. ASER's intended influence pathways depict ASER's "theory of change" embedded in its design (Results for Development Institute, 2015, p. 12). The pathways explain how ASER's architects perceive ASER fostering social betterment, which they describe as an improved education system and better educational practices.

ASER's theory of change includes both an ideal process and the implicit values underlying that process. The theory of change asserts the goals of policy change and grassroots action are achieved via a process of (1) public dialogue at national and state levels and (2) awareness at district and community levels. Values of citizen participation, deliberation, *social accountability* (Christie & Alkin, 2013), and *shared responsibility* (Chouinard, 2013) animate ASER's theory of change. These values provide insight into ASER's design choices—why its approach and methods were chosen in response to the Indian context.

Approach and methods. ASER's approach is participatory and hierarchical—engaging local volunteers and partner organizations but directed from above by state and national staff. ASER's methods are village mapping and surveys, school surveys, and household surveys and learning tests. Integrating aspects of monitoring and policy evaluation with principles of participatory program evaluation, ASER's design demonstrates sensitivity to local contexts; engagement of diverse stakeholders in the evaluation process; and promotion of "local ownership, empowerment, … [and] organizational and individual learning" (Chouinard, 2013, pp. 237–238).[8] ASER's approach and methods cannot be genuinely appraised apart from knowledge of the rural Indian context for which it was created.

ASER's Context

Several things make India a complex evaluation context: its population's size and diversity (e.g., ethnic, caste, religious, and linguistic); its communal inequalities and tensions (Goodnight, 2017b); the remoteness and terrain of its rural districts; and the scale of its civil society and government.

History and politics. ASER was created during a national push to achieve universal primary education under India's commitment to realizing *Education for All*; today, ASER data continues to have relevance under the fourth Sustainable Development Goal on inclusive and quality education for all children. ASER was developed in 2004 by Pratham, India's largest educational non-governmental organization (NGO). Pratham piloted ASER following its discovery that too many children attending Pratham education programs were not learning as expected given their regular primary school attendance. The NGO's leadership wondered how prevalent the issue was beyond the communities in which it worked. Were children struggling to learn nationwide? At the time, Pratham's leadership was also participating in national policy discussions. Pratham leaders cultivated a distinct outlook on the education situation in India with a dual focus on (1) schooling at the grassroots level, and (2) policymaking, governance, and data use at the national level. Pratham saw a need for consistent nationwide data that could be disaggregated to the state- and district-levels and that would provide systematic evidence of children's schooling and learning in rural areas. ASER could serve as a "proof of concept" to government officials that education data could be collected reliably and inexpensively to inform policymaking, budgetary decisions, and school initiatives. In 2008, Pratham gained funding from Google to establish an independent NGO called ASER Centre to lead ASER moving forward.

ASER monitors progress on priorities outlined in India's Right to Education Act using indicators that can prompt widespread discussion about the status of government schools and children's

learning under this policy. ASER also raises vital questions about the logic of the Act's prescriptions (ASER Centre, 2015). ASER Centre leadership hopes quantitative evidence illuminating the Right to Education Act's implementation and success can create public pressure to motivate beneficial planning and policy changes. ASER strives to raise weighty questions about (1) *public direction* (Schwandt & Gates, 2016)—the adequacy of government's education priorities and the logic of the Right to Education Act, and (2) *cause and effect* (Clarke & Dawson, 1999)—the evident relationships between things like private schooling, home factors, and learning (See ASER Centre, 2015, pp. 19–21).

Culture and language. Certain features signal ASER's innovativeness and cultural attentiveness like its use of simple learning tests, its household approach to testing, and its engagement of local volunteers as field data collectors and partner organizations as facilitators (Goodnight, 2017a; Goodnight & Bobde, 2018). These innovations make ASER's massive scope possible (logistically and financially) while strengthening ASER's representativeness and *multicultural validity* (Kirkhart, 1995, 2010). For instance, it is a significant decision to use one-page reading, arithmetic, and English language tests that each assess four levels of competency up to a second-grade proficiency. Compared with pen-and-paper tests, the oral and one-on-one administration of ASER tests by field-trained, local volunteers increases the validity of ASER learning data for children who have limited literacy, limited exposure to educational testing, and whose mother tongues differ from their language of instruction. The tests' simplicity enables thousands of volunteers to assess learning reliably and illiterate parents to decipher the tests' levels and meaning, bolstering awareness about children's learning within rural communities where many adults have limited schooling. Household testing includes children who are out of school and enrolled in various kinds of schools, while facilitating data collection on home factors influencing schooling (Goodnight & Bobde, 2018). Another feature is using local volunteers—typically young pre-service teachers, university students, or NGO workers—as fieldworkers, which helps ASER to transcend vast cultural and linguistic differences and increases the trustworthiness of its data. ASER engages young people in rural education issues and exposes them to social research methods while minimizing ASER's costs (Byker & Banerjee, 2015). Likewise, partner organizations—usually government teacher training institutes, universities, or NGOs—help minimize the financial and logistical hurdles of conducting ASER because they provide personnel and space, and have vital local knowledge (regarding communities' language, culture, and security issues) that makes it possible to execute ASER in districts with difficult terrains, social tensions, and other challenges. While the cultural and linguistic advantages of ASER's design are noteworthy, the safety, health, and fair compensation of ASER volunteers, partners, and staff are weighty concerns warranting further research; investigating these issues may uncover unintended negative consequences of ASER's participatory design.

The Global Influence of ASER's Design

Though ASER's design reflects the Indian context, civil society stakeholders in Global South nations who want to increase accountability and foster awareness about learning in their countries have been adapting ASER's design to their context. ASER has achieved the status of a transnational model with key design features being adopted by similarly-minded NGO and development organizations (e.g., Jàngandoo in Senegal and, Uwezo in Uganda, Kenya, and Tanzania). The features unifying these M&E efforts across countries are (1) testing "basic reading and math competencies"; (2) conducting testing and surveying in households rather than schools; (3) conducting testing "orally and one-on-one"; (4) producing data that are "statistically representative"; (5) executing an evaluation that is independent from government; (6) testing basic skills with a few tasks for all children "regardless of age or grade-level"; (7) engaging citizens in data collection; and (8) timing data to be useful to

a "broad audience, not just authorities and policymakers" (Results for Development Institute, 2015, p. 11). All these features emanating from ASER's design were intentional choices made by its architects during its development. To further support cross-country efforts to employ ASER's citizen-engaged evaluation model (called "citizen-led assessments"), the People's Action for Learning (PAL) Network officially formed in 2015 with financial support from the Hewlett Foundation.[9]

For many global education actors, generating national data on learning is a key goal for monitoring progress on the fourth Sustainable Development Goal. A 2015 report produced by Results for Development Institute found that ASER impacted global discourse and *agenda setting* through (1) increasing focus on learning outcomes as evidence of a serious "learning crisis" and (2) proving "a low-resource model" can effectively generate national learning data (p. 4). Agenda setting is a previously identified form of evaluation influence: a coalitional change mechanism defined as an issue elevated on the agenda of media, government, or the public (Henry & Mark, 2003a). The ASER model continues to garner interest from foreign academics, multilateral institutions, think tanks, and development bodies and influence the global educational agenda. For example, UNESCO and the Brookings Institution together founded the Learning Metrics Task Force (2012–2016), which was co-chaired by ASER architect Rukmini Banerji. The Task Force determined that "measures for globally tracked indicators must be a public good, with tools, documentation and data made freely available" (Learning Metrics Task Force, 2013, p. 12). The ASER citizen-engaged model for evaluating education systems is just that: a solution for producing inexpensive, transparent, and independent large-scale data crucial for closing a "global data gap on learning outcomes" and for ameliorating a crisis of poor-quality education (Learning Metrics Task Force, 2013, p. 9). Increasingly, academics outside India cite ASER data, expressing either concern or support for its model's proliferation (e.g., Alcott & Rose, 2015; Barrett, 2011; Winthrop & Simons, 2013).

ASER's design diffusion. The transnational influence of ASER's design has been incremental and initially unintended. While ASER did begin as a proof of concept with considerable attention paid to its design decisions, it was created for the purpose of persuading the Indian government of the efficacy of its model for examining India's rural education system (Goodnight, 2017a). ASER's global impact was not part of its original purpose (though ASER Centre and Pratham have since encouraged interest and adoption of the model). The American Evaluation Association's definition for evaluation design emphasizes investigating a particular evaluand in a specific context with a clear purpose. Thus, one significant aspect of ASER's influence is that once its evaluation design became a transnational model, its theory and methodological choices became unmoored from the context for which it was created, and its design's impact transcended its original stakeholders and intended audience of its evaluation. Consequently, ASER's *design-based influence* is characterized by the impact of its evaluation model (i.e., its theory, approach, methods, and type of data) in terms of agenda-setting and the development of future evaluation in entirely new contexts. This second design-based influence of ASER relates to the concept of *diffusion* in the evaluation influence literature, which "refers to the adoption of the policy, program, or practice being evaluated in a jurisdiction outside of the one in which the evaluation was conducted" (Henry & Mark, 2003a, p. 305). However, ASER's form of influence is better characterized as *design diffusion* with the thing being adopted not what is under evaluation but rather the model of evaluation. Mark (2011) previously noted the phenomena of diffusion of an evaluation's design, but he indicated that the source is the evaluation process. In ASER's case, the diffusion of its design is not rooted primarily in people's participation in ASER's process. Rather, its design diffusion is linked to the multifaceted activities of a global policy-shaping community, and the ideas and relationships shaping the exchanges between civil society actors in Global South countries. While some actors have traveled to India to observe phases of the ASER evaluation, it was with prior knowledge of ASER's model as an impetus. Transnational interest in being part of

ASER's process is for the purpose of studying its citizen-engaged evaluation design in action and learning how to adapt it for somewhere else.

Why does design diffusion occur? Theorizing from this study of ASER, an evaluation's design can be influential due to its innovativeness in a variety of ways, such as (1) how it addresses common barriers to evaluating a particular evaluand (e.g., rural education), (2) how it overcomes challenges to conducting evaluation in a specific context (or type of contexts), (3) how its theoretical underpinnings or values reshape the purpose of evaluation, and (4) how its approach uniquely involves various stakeholders in the evaluation. Further research on how the ASER model has been adapted in other national contexts and what influence the model is having on different education systems and stakeholders is important to undertake. With its history in the field alongside conversations about knowledge utilization and policymaking (e.g., Ottoson & Hawe, 2009; Rich, 1991; Weiss, 1977), *diffusion theory* provides a foundation for pursuing future RoEI of this model. Drawing on Rogers (2003) diffusion framework, Ashley (2009) highlights key areas that could inform the ASER model's investigation: its adaptability, the social system (i.e., contexts) that facilitates its adoption, communication channels through which it is adopted, the timing of its adoption, and the processes of its adoption. In exploring why ASER was adopted in each context, diffusion theory offers a set of factors for analysis: an innovation's *relative advantage*, *compatibility*, *complexity*, *trialability* and *observability* (Ashley, 2009).

Conclusion

Several insights emerge from using this RoEI conceptual framework to analyze the ASER case:

- The source dimension previously conceived in evaluation influence theories is incomplete. Beyond an evaluation's process and results, there is an important source of influence: an evaluation's design.
- Stemming from ASER's design-based influence is a form of evaluation consequence—design diffusion—where the innovation being adopted is the model of evaluation itself.
- Focusing on design as a source of influence facilitates analysis of evaluation context and an evaluation's antecedents. Studying an evaluation design initiates a RoEI researcher's engagement with issues of politics, culture, history, and language present in the evaluation setting (to which the evaluation's theory, approach, and methods may or may not be attuned). Paying attention to design in a study of evaluation influence takes into account Hall's (2004) points that the antecedents of evaluation matter and whether an evaluation ultimately has influence is rooted in its responsiveness to its context (inclusive of its design choices).
- While the dimensions of this RoEI conceptual framework are instructive in contemplating multiple aspects of an evaluation's influence, it is hard to disentangle discussion of the dimensions from one another. It was not possible to present the ASER case according to all dimensions independently because for instance, talking about intention (via the intended influence pathways) required discussion of motivation, time, level, and form. In fact, the dimensions are inseparable—co-constructing how to analyze and express an evaluation's consequences.
- The complexity of studying evaluation influence is not lessened by using this RoEI conceptual framework. Moreover, adequately studying all dimensions of evaluation influence in one study is not possible (Mark, 2011). For example, this study includes only a brief examination of the context dimension, as it pertains to ASER's design. In prioritizing the source and intention dimensions, the study is not a deep investigation of the effects of language, culture, politics, or history on ASER's influence. Additional analyses addressing missed opportunities for increasing ASER's contextual and cultural responsiveness could improve its influence in the future.

- Values are omnipresent in evaluation (Greene, 1997), but they are not represented in this RoEI conceptual framework. Values seemingly relate to the motivation and context dimensions, but further research can make better sense of these relationships. In this article, examining the theory underlying ASER's design involved considering its architects' values in developing it. An analysis of values throughout an evaluation's story includes understanding their origins—not just contextual (e.g., cultural and political)—but also philosophical and moral. The cultural, political, philosophical, and moral origins of an evaluation's values lie not only within the evaluation's context and communities, but also within the evaluator. Therefore, from its design to its results, the role of values in determining an evaluation's influence is worthy of future investigation. Equally important is considering the role of values in forming the RoEI researcher's lens on what is influential.

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 would like to acknolwedge the United States Department of State and the Government of India for funding the Fulbright-Nehru student research grant that supported this study's data collection.

ORCID iD

Melissa Rae Goodnight https://orcid.org/0000-0003-1106-6665

Notes

1. The ethnographic case study of ASER resulted in roughly 90 interviews along with hundreds of documents collected and hours of field observation. See Goodnight, 2017a for further details about its process and methods.
2. India is comprised of 28 states and 8 union territories. ASER 2014 included all rural districts across states resulting in over 560 of India's 593 districts surveyed (ASER Centre, 2015, p. 63).
3. Prior to the field's evaluation influence literature, scholars were mapping pathways to depict different ideas regarding evaluation use. Johnson's (1998) highlights the frequency and legacy of visual depictions of utilization in the field; he examines several evaluation theorists' implicit and explicit *process-models* (i.e., visual presentations of variables and processes of evaluation utilization), which he uses to construct his meta-model of evaluation use. Included in Johnson's analysis of process-models is Greene's (1988) "grounded-theory model" of evaluation utilization (p. 97); her diagram could be interpreted as a precursor to those of intended influence pathways. Based on empirical data from two evaluations (using the same evaluation model), Greene's diagram includes observed effects and design elements to describe the phases of a participatory evaluation's process through its utilization. Johnson's (1998) meta-model is a theoretical one stemming from a synthesis of others' process-models rather than an empirical analysis of an evaluation.
4. This case study revealed that findings from ASER's evaluation are not routinely shared with (i.e., systematically disseminated in) sampled villages as part of ASER's evaluation model. Thus, grassroots action from these communities is conceived as stemming from their participation in ASER's process.
5. Gildemyn (2014) highlights that monitoring and evaluation efforts led by civil society organizations in the Global South (i.e., "CSO-led M&E") often seek to stimulate public dialogue and government pressure via media coverage of findings.
6. Gildemyn (2014) provides a useful exemplar by employing case study data of a Ghanaian monitoring and evaluation effort to explore the M&E's observed influence pathways.
7. The definition was included in the American Evaluation Association's 2016 conference theme that focused on evaluation design (held in Atlanta, Georgia, October 24-29, 2016).

8. Chouinard (2013) includes "use of findings" as another aim of participatory evaluation. ASER's design emphasizes process use and influence (not results use) with master trainers, volunteers, community members, and families. However, the leaders of ASER Centre have experimented with different ways to share data with partner organizations and some villages. They have not only expressed this need but also emphasized the resource intensity of sharing findings and supporting community deliberation and action. At the time of this study's data collection, ASER Centre secured some additional funding to pilot different strategies for community and partner strategizing based on ASER findings.
9. The PAL Network formalized organic "south-south partnerships" that started forming in 2006 between education workers and organizations interested in the ASER model. See PAL Network's website for further information: https://palnetwork.org/.

References

Alcott, B., & Rose, P. (2015). Schools and learning in rural India and Pakistan: Who goes where, and how much are they learning? *Prospects*, *45*(3), 345–363.

Alkin, M. (2013). *Evaluation roots: A wider perspective of theorists' views and influences (second edition)*. Thousand Oaks: Sage.

Alkin, M., & Taut, S. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*, *29*, 1–12.

Alkin, M. C., & Christie, C. A. (2019). Theorists' models in action: A second look. In C. A. Christie & M. C. Alkin (Eds.), *Theorists' models in action: A second Look. New directions for evaluation* (Vol. 163, pp. 11–18). Thousand Oaks: Sage.

Alkin, M. C., & King, J. A. (2016). The historical development of evaluation use. *American Journal of Evaluation*, *37*, 568–579.

Alkin, M. C., & King, J. A. (2017). Definitions and factors associated with evaluation use and misuse. *American Journal of Evaluation*, *38*, 434–450.

ASER Centre (2014). *Quality control framework.* New Delhi: ASER Centre.

ASER Centre (2015). *Annual Status of Education Report* (Rural) 2014. New Delhi, India.

Ashley, S. R. (2009). Innovation diffusion: Implications for evaluation. In J. M. Ottoson & P. Hawe (Eds.), Knowledge utilization, diffusion, implementation, transfer, and translation: Implications for evaluation. *New Directions for Evaluation*, 124, 35–45.

Barrett, A. (2011). An education millennium development goal for quality: Complexity and democracy. *Compare*, *41*(1), 145–148.

Byker, E., & Banerjee, A. (2015). Evidence for action: Translating field research into a large scale assessment. *Current Issues in Comparative Education Journal*, *18*, 42–53.

Chilisa, B., Major, T. E., Gaotlhobogwe, M., & Mokgolodi, H. (2016). Decolonizing and indigenizing evaluation practice in Africa: Toward african relational evaluation approaches. *Canadian Journal of Program Evaluation*, *30*(3), 313–328.

Chouinard, J. (2013). The case for participatory evaluation in an era of accountability. *American Journal of Evaluation*, *34*(2), 237–253.

Chouinard, J., & Cousins, J. (2009). A review and synthesis of current research on cross-cultural evaluation. *American Journal of Evaluation*, *30*(4), 457–494.

Christie, C., & Alkin, M. (2013). An evaluation theory tree. In M. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences, second edition* (pp. 11–57). Thousand Oaks: Sage.

Clarke, A., & Dawson, R. (1999). *Evaluation research: An introduction to principles, methods, and practice*. SAGE.

Coryn, C., Ozeki, S., Wilson, L., Greenman, G., Schröter, D., & Hobson, K., … A. Vo (2016). Does research on evaluation matter? Findings from a survey of American evaluation association members and prominent evaluation theorists and scholars. *American Journal of Evaluation*, *37*(2), 159–173.

Cronbach, L. (1982). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.

Gildemyn, M. (2014). Understanding the influence of independent civil society monitoring and evaluation at the district level. *American Journal of Evaluation*, *35*(4), 507–524.

Goodnight, M., & Bobde, S. (2018). Missing children in educational research: Investigating school-based versus household-based assessments in India. *Comparative Education*, *54*(2), 225–249.

Goodnight, M. (2017b). Critical race theory in India: Theory translation and the analysis of social identities and discrimination in Indian schooling. *Compare: A Journal of Comparative and International Education*, *47*(5), 665–683.

Goodnight, M. (2017a). *Learning to "measure what affects their lives": Ethnography of a citizen-engaged evaluation of primary education in India* (10599943). [Doctoral Dissertation, University of California Los Angeles]. ProQuest.

Government of India (2014). Chapter 13: Human development, economic survey 2013–2014 (pp. 230–254). New Delhi: Ministry of Finance.

Greene, J. C. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, *12*(2), 91–116.

Greene, J. C. (1997). Evaluation as advocacy. *Evaluation Practice*, *18*(1), 25–35.

Hall, M. (2004, November). *Persuasive language, responsive design: Towards broader evaluation utilization*. In K. Kirkhart (Chair), What is the nature of evaluation use/ influence? Panel at the American Evaluation Association conference, Atlanta, GA.

Henry, G., & Mark, M. (2003b). Toward an agenda for research on evaluation. In C. M. Christie (Ed.), The Practice-Theory Relationship in Evaluation. *New Directions for Evaluation*, 97, 69–80.

Henry, G., & Mark, M. (2003a). Beyond use: Understanding evaluation's Influence on attitudes and actions. *The American Journal of Evaluation*, *24*(3), 293–314.

Hood, S. (2004). A journey to understand the role of culture in program evaluation: Snapshots and personal reflections of one african American evaluator. *New Directions for Evaluation*, *102*, 21–37.

Hopson, R. (2009). Reclaiming knowledge at the margins: Culturally responsive evaluation in the current evaluation moment. In K. Ryan & J. Cousins (Eds.), *The SAGE international handbook of educational evaluation* (pp. 429–446). Thousand Oaks: Sage.

House, E., & Howe, K. (2000). Deliberative democratic evaluation. *New Directions for Evaluation*, *85*(3), 3–12.

Johnson, R. B. (1998). Toward a theoretical model of evaluation utilization. *Evaluation and Program Planning*, *21*(1), 93–110.

King, J. (2016). Taking what action for change? Dare evaluation build a new social order? *Visitor Studies*, *19*(1), 3–11.

King, J. A., & Alkin, M. C. (2019). The centrality of use: Theories of evaluation use and influence and thoughts on the first 50 years of use research. *American Journal of Evaluation*, *40*(3), 431–458.

Kirkhart, K. (1995). Seeking multicultural validity: A postcard from the road. *Evaluation Practice*, *16*(1), 1–12.

Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. In V. J. Caracelli & H. Preskill (Eds.), The Expanding Scope of evaluation Use, *New Directions for Evaluation*, 88, 5–23.

Kirkhart, K. (2010). Eyes on the prize: Multicultural validity and evaluation theory. *American Journal of Evaluation*, *31*(3), 400–413.

Kirkhart, K. E. (2011). Culture and influence in multisite evaluation. In J. A. King & F. Lawrenz (Eds.), Multisite evaluation practice: Lessons and reflections from four cases. *New Directions for Evaluation*, 129, 73–85.

Learning Metrics Task Force (2013). *Toward universal learning: Recommendations from the learning metrics task force UIS/2013/ED/TD/07*. Washington, D.C.: UNESCO Institute for Statistics and Center for Universal Education at Brookings Institution.

Mark, M. (2011). Toward better research on—and thinking about—evaluation influence, especially in multisite evaluations. In J. A. King & F. Lawrenz (Eds.), *Multisite evaluation practice: Lessons and reflections from four cases. New directions for evaluation* (Vol. 129, pp. 107–119).

Mark, M., & Henry, G. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, *10*(1), 35–57.

Nevo, D. (1982). The International Context for Research on Evaluation. *American Journal of Evaluation*, 3(4), 73–75. https://doi.org/10.1177/109821408200300418

Nunneley, R.Jr., King, J., Johnson, K., & Pejsa, L. (2015). The value of clear thinking about evaluation theory: The example of use and influence. In C. Christie & A. Vo (Eds.), *Evaluation Use and decision making in society: A tribute to marvin C. Alkin* (pp. 53–71). Charlotte: Information Age Publishing.

Ofir, Z., & Kumar, A. (2013). Evaluation in developing countries. In S. I. Donaldson, T. Azzam, & R. F. Conner (Eds.), *Emerging practices in international development evaluation* (pp. 11–24). Charlotte, NC: Information Age Publishing.

Ottoson, J. M., & Hawe, P. (2009). Knowledge utilization, diffusion, implementation, transfer, and translation: Implications for evaluation. *New Directions for Evaluation*, *124*.

Patton, M. Q. (2020). Evaluation use theory, practice, and future research: Reflections on the alkin and king AJE series. *American Journal of Evaluation*, *41*(4), 581–602.

Project Team, T. O. H. (2006). The oral history of evaluation, part 4: The professional evolution of carol H. Weiss. *American Journal of Evaluation*, *27*(4), 475–484.

Results for Development Institute. (2015). Bringing Learning to Light: The Role of Citizen-led Assessments in Shifting the Education Agenda. Washington, D.C.

Rich, R. F. (1991). Knowledge creation, diffusion, and utilization: Perspectives of the founding editor of knowledge. *Knowledge*, *12*(3), 319–337.

Rogers, E. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.

Schwandt, T., & Gates, E. (2016). What can evaluation do? An agenda for evaluation in service of an equitable society. In S. Donaldson & R. Picciotto (Eds.), *Evaluation for an equitable society* (pp. 67–82). Charlotte: Information Age Publishing.

Weiss, C. (1977). *Using social research in public policy making*. Lexington, Massachusetts: Lexington Books.

Weiss, C. (1999). The interface between evaluation and public policy. *Evaluation*, *5*, 468–486.

Winthrop, R., & Simons, K. (2013). Can international large-scale assessments inform a global learning goal? Insights from the learning metrics task force. *Research in Comparative and International Education*, *8*(3), 279–295.