Title

Learning to "Measure What Affects their Lives": Ethnography of a Citizen-Engaged Evaluation of Primary Education in India

Author

Goodnight, Melissa Rae

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Learning to "Measure What Affects their Lives": Ethnography of a Citizen-Engaged Evaluation of Primary Education in India

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Education

by

Melissa Rae Goodnight

2017

ABSTRACT OF THE DISSERTATION

Learning to "Measure What Affects their Lives":

Ethnography of a Citizen-Engaged Evaluation

of Primary Education in India

by

Melissa Rae Goodnight

Doctor of Philosophy in Education

University of California, Los Angeles, 2017

Professor Christina A. Christie, Co-chair

Professor Edith S. Omwami, Co-chair

The *Annual Status of Education Report (ASER)* evaluates primary education and children's learning across rural districts of India. ASER is a *citizen-engaged evaluation*, relying on the participation of roughly 25,000 volunteers and over 500 partner organizations every year. ASER volunteers conduct village asset mapping, government school surveys, and household-based surveys and learning tests, which provide the only annual data on Indian children's basic learning levels. The design and process of ASER directed by its central evaluation question: "Are our children learning?" Because of its household-based approach, capacity-building components, and inclusion of volunteers and partners, ASER represents a unique model for social systems evaluation.

My ten-month ethnographic study is the first in-depth research on ASER's design and the only study to date to analyze ASER as an evaluation model. To investigate ASER's design and influence in India and globally, I utilized document analysis, field observations in three states, and drew upon interviews with ASER's architects, funders, staff, volunteers, and partners. Given its geopolitical significance, its population's diversity, and the scale of its civil society and government, India is an important context for research on evaluation—especially with respect to expanding current evaluation theories, which mostly derive from Western contexts and smaller scale evaluations.

My findings explain 1) why ASER was developed, 2) how ASER's pioneering methodology was devised, 3) what implications ASER has for India's education system and society, and 4) how ASER's model for evaluation has gained global *design influence*. There are several ways to understand ASER's significance—the impact of ASER's results on educational policy and governance in India, the influence of ASER participation as a social process involving ordinary people and local institutions, and the popularity of ASER's model for evaluation, which has been borrowed between countries and advocated by multilateral organizations.

Utilizing a three-paper structure, this dissertation illuminates ASER's contribution to a culture of evaluation in India by examining ASER's development, design, and process. This study theorizes the connection between evaluation, democratic governance, and citizen participation in the improvement of India's rural primary education system and specifically, of children's learning across the country.

Keywords: monitoring and evaluation, India, ethnography, narrative inquiry, evaluation theory, *design influence*, education reform

This dissertation of Melissa Rae Goodnight is approved.

Akhil Gupta

José Felipe Martínez

Christina A. Christie, Committee Co-chair

Edith S. Omwami, Committee Co-chair

University of California, Los Angeles

2017

DEDICATION

To my family, especially my parents, Michael and Elaine Goodnight.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

Like the *Annual Status of Education Report* (ASER), a dissertation is a collective accomplishment. I could not have completed this study of ASER without the immense encouragement and support that I have received from many people.

I would like to express my deep appreciation for my dissertation co-chairs Dr. Edith Mukudi Omwami and Dr. Tina Christie and my dissertation committee members, Dr. José Felipe Martínez and Dr. Akhil Gupta.

During my first year at UCLA, Edith told me she chose me as her doctoral advisee not because she wanted an academic clone but because she believed that I deserved a chance to do meaningful work. Since then, she has demonstrated in every interaction an unwavering belief in my capacity and a genuine willingness to let me evolve in my skills and identity as a scholar—stating things like "Melissa, you must discover and do your own research for the rest of your career, not mine." I have found that Edith uncannily knows what to say when I am at a loss: more recently, when I was feeling discouraged, she explained: "Melissa, a dissertation is just the beginning of a long conversation." Edith has supported me with ongoing compassion, humor, and wisdom. This dissertation reflects her intellectual influence especially in its tackling of global educational policy and institutions from an equity and human rights framing.

I am grateful for Tina's mentorship, guidance, generosity, and resourcefulness. Recounting a history jam-packed with compelling characters and interesting debates (see *Evaluation Roots*, 2013), Tina introduced me to the field of evaluation and to evaluation theory. Tina has since guided me and thus patiently listened to me theorize about my dissertation project on evaluation. She has wholeheartedly supported my research on evaluation and used all of her resources and professional connections to facilitate it. This dissertation would quite simply

neither exist without her outstanding teaching and mentorship, nor without her substantive research on evaluation theory and its significant impact on my understanding of the field.

I am thankful for the several ways in which Felipe has influenced my scholarly identity and nurtured my intellectual development, particularly in the areas of research design and measurement. Felipe's teaching has had a direct and lasting impact on my ability to think critically and has fundamentally altered how I perceive quantitative and mixed methods research. I have vetted a lot of half-formed ideas in his research apprenticeship class to which I have always received the right balance of encouragement and skeptical eyebrow raising. Ultimately, the rigor of his teaching paired with his guidance and encouragement have fostered my confidence in my capacity to engage in these methodological spheres. I conclude this dissertation project as a methodological generalist but with an emergent (perhaps, peculiar) specialization in the ethnography of big data for which I can only thank him.

I could not have been more thankful to discover Akhil's research, *Red Tape: Bureaucracy, Structural Violence, and Poverty in India* (2012), and subsequently gain his insight on my dissertation. I have not only benefited from Akhil's wide-ranging expertise on the administration of social programs and their relationship to rural, impoverished, and marginalized communities within rural India, but I have also grown in my understanding of how to recognize the nuanced and contradictory effects of such programs as a result of his scholarship and its use of theory. Reading Akhil's *Red Tape* taught me substantive lessons on how to grapple with complexity in one's findings and writing.

In addition to my dissertation committee, I owe special thanks to my friends, mentors, and colleagues from Ambedkar University Delhi, Pratham Education Foundation, and ASER Centre, who have made my research possible and whose dynamic and heartfelt work is the

central focus of this dissertation. Over the course of the past five years I have spent conducting research across the country, it is the generosity of my colleagues and mentors in India who have made my scholarship both successful and meaningful. Not only have they shared their professional knowledge and experiences with me but also opened up their homes to me; welcomed me like family into their everyday lives and festivals; and, consistently helped me to navigate a vast nation comprised of many distinct places, cultures, and languages.

Several of my mentors and colleagues in India have become both cherished teachers and friends. This dissertation is simultaneously a reflection of their extraordinary kindness to me and of their enduring commitment to education, social research, and to their country. I particularly want to thank Dr. Suman Bhattacharjea, Dr. Wilima Wadhwa, Dr. Madav Chavan, and Dr. Rukmini Banerjee of ASER Centre and Pratham for their openness, guidance, and unfailing support for this study; Dr. Venita Kaul and Dr. Gunjan Sharma of Ambedkar University Delhi for their insightful mentorship; Ruchi Aggarwal and Surbhi Batra Mahandru for being my first colleagues and oldest friends in India and for teaching me about rural education in India and conducting fieldwork in villages; the Batra family for being my loving and supportive Delhi-based family; the Banerjee family for their extraordinary kindness and encouragement; Ashok Mutum and the Mutum family for their generous care and support; Savitri Bobde for her mentorship, friendship, and sharing her comprehensive knowledge; Remy Hans for offering me remarkable help with fieldwork and translation in Rajasthan as well as for her friendship; and Karthika Anthony and Nandita Banerjee for their dear friendship, invaluable advice, and fabulous senses of humor and adventure. Finally, I want to acknowledge the Manipur, Tamil Nadu, and Rajasthan state teams with whom I was imbedded while doing fieldwork and whose members were absolutely integral to my dissertation (but whom I do not name individually to

protect their confidentiality). I cannot easily express the depth of my gratitude for their inordinate thoughtfulness, generosity, and assistance in helping me collect data as we traveled together by train, bus, jeep, car, and motorbike—sometimes, for weeks at a time. It was not easy for them to assume the responsibility of my welfare amidst their own responsibilities for ASER, but they did so in addition to fully dedicating themselves to helping me collect the best data possible for my project.

My dissertation research was supported through a Fulbright-Nehru Student Research Grant (funded by the U.S. Department of State and the Government of India) and a Critical Language Enhancement Award (CLEA) for Hindi and Urdu study. This research was also funded by University of California Los Angeles through the UCLA Affiliates Scholarship and UCLA Dissertation Year Fellowship. The Fulbright-Nehru and CLEA awards were administered by the United States-India Educational Foundation (USIEF), and I want to thank USIEF staff members—especially, Neeraj Goswami, Pavitra Soram, and Diya Dutt–for supporting me during my language training and fieldwork for this dissertation. Through USIEF training sessions and conferences in India, I gained the crucial support and guidance of Fulbright-funded, U.S. professors conducting research in India who became my de-facto advisory committee abroad: Dr. Christine Finnan of the College of Charleston and Dr. Steve Hite and Dr. Julie Hite of Brigham Young University as well as Dr. Tamo Chattopadhay (formerly from India's National University of Educational Planning and Administration and now with Bard College's Institute for International Liberal Education). I am extremely grateful to them for their timely and free-flowing advice. Also, Dr. Gunjan Sharma of Ambedkar University Delhi thoughtfully supervised my Fulbright research and offered vital policy and practice insights on the Indian education system.

Dissertation research is the capstone of a long academic journey. I am grateful to all of the UCLA faculty and staff members, colleagues, friends, and advisors who have helped me along mine. In particular, I want to thank the following people: Harmeet Singh, Ryan Lebre, and Amy Gershon of the Graduate School of Education and Information Studies (GSE&IS) have kindly helped me navigate the many stages of my Ph.D. program. Dr. Stephen Commins of the Department of Urban Planning has guided me through many phases of my research in India and mentored me on issues of international development and policy. Dr. Gyanam Mahajan of the Department of Asian Languages and Cultures coached me in several of the linguistic and cultural foundations of India that were essential to my dissertation fieldwork. Dr. Gary Rhodes thoughtfully advised me on the Fulbright grant program that ultimately funded this dissertation, and offered guidance on living abroad and conducting research in India. Dr. Sandra Harding of GSE&IS was an early advocate of my research and writing; our conversations regarding the philosophical and political foundations of my work enhanced my understanding of my positionality in this project. Dr. Marv Alkin of GSE&IS enthusiastically advised me on the argument of my final paper in this dissertation and shared with me his encyclopedic knowledge of evaluation theory—my theorizing has enormously benefited from his counsel and his profound work in the field of evaluation. Dr. Mike Rose was an imaginative and kind home-stretch coach, who through his stellar example taught me a great deal about the psychology and art of writing and about how to teach and support others. Dr. Kathryn Anderson-Leavitt has time and again read and insightfully commented on not only my analytic memos and early writings of my dissertation, but also entire chapters and manuscripts for publication. Her wise advice has been indispensable. I am genuinely awed by her profound and tireless commitment to students' development as researchers and writers.

My colleagues and friends from GSE&IS and the wider comparative education community have mentored me and inspired me with their brilliant thinking and deep dedication to their research, teaching, and service. I especially want to acknowledge Tatevik Achabahian, Hughlin Boyd, Dr. Sayil Camacho, Dr. Ryan Donaghy, Dr. Jason Dorio, Kareem Elzein, Michelle Gaston, Shey Khaksar, Dr. Nafees Khan, Dr. Jeeyoung Lee, Nelli Martirosyan, Dr. Laura McMullin, Liz Perez-LoPresti, Janelle Rahyns, Kevin Schaaf, Dr. Victor Soler, Dr. Rolf Straubhaar, Jayashri Srinivasan, Talia Stol, Mariana Barragan Torres, Dr. Winmar Way, Hui Xie, and Taiko Yusa (who visited me during my fieldwork and eased my homesickness with her unshakable cheeriness). Thank you all for supporting my scholarship and nurturing my development in countless ways.

My dissertation has been substantially strengthened by the conscientious, creative, and clever editorial advice of my writing group. For the past year and a half, Dr. Kathy Carbone, Dr. Audrey Harris, and Nathan Kish have been my weekly advisors and savvy writing coaches. "Forsooth," I owe so much of my productivity and perseverance to their genius, humor, and support. Alas, I only wish that I were wittier and could figure out a way to stealthily include "the archives," "Dr. Sharma," and something about "feminism and Borges" into this text.

As mentioned previously, academic journeys are long ones. Dr. Karen Monkman was my advisor in my master's program at DePaul University and the person who first introduced me to comparative and international education and qualitative research. Several concepts that ultimately led me to my dissertation on ASER began as kernels of curiosity in Karen's classes. Over the past ten years, Karen has been a powerful example of an educator who is committed to her students and to supporting the public good and social justice through her scholarship.

Finally, I want to acknowledge the heartfelt support I have received from my close

friends and family throughout my Ph.D. program and dissertation process. Special thanks to Dr. Sybil Madison-Boyd, Lauren Schmied, Dr. Shubhra Misra, Kris Sellman-Johnson, Osa Wallander, Sang Lee, Maegen Rose, Aflamu Johnson, Jon-Barrett Ingels, Casey Hirshfield, and Daria Linvill. I am also grateful to Chris and Chrissy Burroughs for their championing of my work in India and to Gary and Laurie Edemann, Harry Orrill, Kathy Hughes, the McCann family, the Martinez family, and the Naik family for their constant encouragement and faith in me. My partner, Amol Naik, has been an amazing coach, advisor, editor, sounding board, and friend throughout my doctoral program. He frequently stayed up odd hours talking with me during my months of dissertation fieldwork in India. Amol (joined in equal measure by my parents) worried from afar over my safety, my happiness, and my data. Even when I was not across the world (but only across the country in Los Angeles), Amol was still losing sleep over editing my lengthy proposals and papers while trying to manage his own demanding studies. In short, my Ph.D. has cost him rest, vacations, a partner who lives in the same city, and my attention for the better part of six years—Amol has nevertheless celebrated every one of my milestones with pride. My doctoral program necessitated his sacrifices and understanding and also that of my family—including my parents, Phil, Nancy, Savannah, and Josh—for which I cannot adequately express my gratitude. I am most thankful for the ongoing guidance, love, and support of my parents, Mike and Elaine Goodnight, who were my first teachers and to whose brilliance, diligence, and care, I owe so much of my success and happiness. I share this achievement with them.

VITA

EDUCATION

2017 Graduate Certificate in Writing Pedagogy
University of California Los Angeles, Los Angeles, CA

2010 M.A. in Education
DePaul University, Chicago, IL

2001 B.F.A. in Theatre Performance
Chapman University, Orange, CA

FUNDING

2016 UCLA Dissertation Year Fellowship
University of California Los Angeles, Los Angeles, CA

2014 Fulbright-Nehru Student Research Grant
Fulbright Scholarship Board, U.S. Department of State & Government of India

2014 Critical Language Enhancement Award (Hindi)
Fulbright Scholarship Board, U.S. Department of State

AWARDS AND HONORS

2016 Research on Evaluation Student Award
American Evaluation Association

2015 Distinguished Teaching Assistant/Special Reader Award
Graduate School of Education & Information Studies, UCLA

2013 Prize in Memory of Leigh Burstein
Graduate School of Education & Information Studies, UCLA

2013 Departmental Honors, Doctoral Qualifying Exams
Social Science & Comparative Education (SSCE) Division, UCLA

PROFESSIONAL EXPERIENCE

2014, 2015 Course Instructor, English Composition
University of California Los Angeles, Writing Programs

2013 Assistant Instructor, Comparative Education
University of California Los Angeles, Department of Education

2012-2014 Lead Editor & Teacher Evaluation Special Section Editor
InterActions: UCLA Journal of Education & Information Studies

PUBLICATIONS

Goodnight, M. (forthcoming). The Language-Related Academic Self-Confidence of Non-citizens in U.S. Colleges. *Journal of College Student Development.*

Goodnight, M. (2017). 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.* [Published online January 18, 2017: http://www.tandfonline.com/doi/full/10.1080/03057925.2016.1266926]

Goodnight, M. and Acker, A. (2013). Critical Thinking as an Everyday Practice: A Discussion with Sandra Harding about the History of *InterActions*, Interdisciplinary Scholarship, and her New Book. *InterActions: UCLA Journal of Education and Information Studies*, 9(1).

CONFERENCE PRESENTATIONS

Goodnight, M. (2017, April). *The Use of Narrative Inquiry in Research on Evaluation.* Paper presented at the annual meeting of American Educational Research Association, San Antonio, Texas.

Goodnight, M. (2016, October). *Design Influence: Investigating the Global Momentum Behind India's ASER model of Evaluation.* Paper presented at the annual meeting of American Evaluation Association, Atlanta, GA.

Goodnight, M. (2016, April). *Citizen-led Evaluation of Education in Rural India: Partner Organizations' and Volunteers' Participation in ASER.* Paper presented at the annual meeting of Center for Culturally Responsive Evaluation and Assessment, Chicago, IL.

Goodnight, M. & Bobde, S. (2016, April). *School-based Assessment in Rural India: Representation and the Pursuit of Education for All.* Roundtable presented at the annual meeting of American Educational Research Association, Washington, D.C.

Goodnight, M. (2014, April). *The Language-Related Academic Self-Confidence of Non-citizens in U.S. Colleges*. Paper presented at the annual meeting of American Educational Research Association, Philadelphia, PA.

Goodnight, M. (2013, March). *Researching People Like You/Unlike You: Dichotomies and their Meaning in Feminist Postcolonial Fieldwork in Education and Development.* Paper presented at annual meeting of Comparative and International Education Society, New Orleans, LA.

Goodnight, M. (2013, March). *Critical Race Theory in India? Struggling with theory translation in an analysis of social identities and discrimination in Indian schooling.* Paper presented at annual meeting of Comparative and International Education Society, New Orleans, LA.

INTRODUCTION

The *Annual Status of Education Report (ASER)* evaluates primary education and children's learning across rural districts of India. ASER is a *citizen-engaged evaluation*,[1] relying on the participation of roughly 25,000 volunteers and over 500 partner organizations every year. The evaluation is conducted entirely independent of the Government of India (GOI) by the Delhi-based non-governmental organization, ASER Centre. ASER volunteers conduct village asset mapping, government school surveys, and household-based surveys and learning tests, which provide the only annual data on Indian children's basic learning levels. The central question determining ASER's evaluation design is "Are our children learning?"[2] ASER represents a unique model for the evaluation of social systems (e.g., education). My 10-month ethnographic study is the first in-depth research on ASER's design and the only study to date to analyze ASER as an evaluation.[3] To investigate ASER's design and the influence of its model within India and globally, I conducted document analysis, field observations in three states, and interviews with ASER's architects, funders, staff, volunteers, and partners.

Given its geopolitical significance, its population's diversity, and the scale of its civil society and government, India is an important context for research on evaluation—especially

[1] A *citizen-engaged evaluation* is a model of large-scale systems evaluation, which relies upon the participation of partner organizations and volunteers organized under the principle that ordinary citizens have rights within a democratic society to independently evaluate their public systems.

[2] When asked "What is the purpose of ASER? What is it trying to do?" ASER Centre leadership and staff members responded that ASER is trying to find out, "Are our children learning?" This question is now the recurring title of the educational reports published by UWEZO, a non-governmental organization that has adapted the ASER model and is conducting its version of ASER in Uganda, Tanzania, and Kenya.

[3] Other research conducted on ASER has been shorter term or more limited in scope. Examples are articles about ASER's evolution written by ASER Centre leadership (e.g., Banerji, Bhattacharjea, & Wadhwa, 2013), methodological papers (e.g., ASER Centre, n.d.), and limited references to or explanations of ASER in academic papers (e.g., Barrett, 2011; Pritchett, 2013). ASER along with translations of its model in other Global South countries was the focus of a recent report on "citizen-led assessments" authored by Results for Development Institute (2015).

with respect to expanding current evaluation theories, which mostly derive from Western contexts and smaller scale evaluations. My findings explain why ASER was developed, how its pioneering methodology was devised, how its process relates to public awareness regarding primary education, how its results are intended to prompt policy change, what implications ASER has for the field of evaluation, and why its model for evaluation has gained *design influence*.[4] There are multiple ways to comprehend ASER's significance—the impact of ASER's results on educational policy and governance in India, the influence of participating in ASER on ordinary people and local institutions, and the increasing popularity of ASER's model for evaluation with other countries and multilateral organizations. This study illuminates ASER's contribution to a "measurement culture" in India and theorizes the connection between evaluation, democracy, and social betterment.

Dissertation Structure

The dissertation follows a multi-paper structure to enable the crafting of separate but related arguments. Each of the three papers addresses different aspects of ASER's design—the development of its methodology and goals, the phases of its citizen-engaged process, and the influence of its design globally.

Paper 1. "We Went in Favor of Rebellion": A Narrative Inquiry into the Design of India's Annual Status of Education Report

ASER is groundbreaking in its scale, cost effectiveness, simple design, and "citizen-led" approach. The non-governmental organization, Pratham, created ASER in response to children's learning issues and as a "proof of concept" to Indian policymakers. ASER's origins in grassroots programming and policymaking distinctly shape its methodology and goals. Describing its

[4] *Design influence* is the influence of an evaluation's design (its theory, approach, methods, process, and type of data) as an innovative model for the development and conduct of further (future) evaluations.

development amid large policy shifts in India's school system, the narratives of ASER's three main architects explain how ASER forged a path complementary to government efforts while promoting citizens' rights to independently evaluate learning. The secondary argument of this paper is about the use of ethnography and narrative inquiry as methodological techniques that are beneficial for studying the history behind and values underlying monitoring and evaluation (M&E) efforts.

Paper 2. Theorizing from the Ground: Ethnography of a Citizen-Engaged Evaluation Process in India

ASER is not only a large-scale technical endeavor but also a far-reaching social process that promotes public awareness, social accountability, shared responsibility, and evaluative thinking. One can analyze ASER as a unique form of evaluation—a model for *citizen-engaged evaluation*—that investigates India's progress in universalizing primary education by innovatively engaging parents, teachers, volunteers, and members of partner organizations in its data collection process. This paper draws upon the academic literature on evaluation in order to examine aspects of ASER's model and process as well as explore its potential as "a kind of social conscience" (Schwandt & Gates, 2016, p. 67). Furthermore, because ASER is an evaluation developed outside of the Global North, this paper posits how theorizing about ASER from the ground may enrich the field's literature and expand the appropriateness of evaluation concepts for non-Western contexts.

Paper 3. *Design Influence*: The Source of an Evaluation's Power Beyond Process and Results.

Through exploring concepts of *evaluation influence* (Kirkhart, 2000), this paper addresses ASER's global *design influence* as a transnational evaluation model. In the educational monitoring and evaluation sphere (M&E), ASER has garnered international attention because of

its groundbreaking scale, cost effectiveness, simple design, and participatory approach. The influence of ASER's design expands current theorizing about evaluation influence that associates the source of influence either with the evaluation process or results. These findings from ASER indicate that an evaluation's design can have influence as a model for evaluation and thus, has impact beyond its original context and intended group of stakeholders.

Study Design and Data Collection

This dissertation study included extensive fieldwork in Delhi, Rajasthan, Manipur, and Tamil Nadu, resulting in hundreds of hours of observation and over 90 semi-structured interviews. I conducted interviews with government officials, Indian academics and educational experts, ASER's architects, central and state staff members, partner organizations and volunteers. I observed phases of the ASER process including trainings, village mapping, testing in households, surveying in schools, and dissemination of data. The study's theoretical framework is a synthesis of evaluation and comparative education concepts: *vertical case study methodology* from comparative education connects micro-level qualitative studies of an educational phenomenon with macro-level influences (Bartlett & Vavrus, 2009), and evaluation theories, particularly those focused on equity, social accountability, cultural responsiveness, democratic deliberation, and participation articulate the relationship between evaluation, social betterment, and democratic social systems (e.g., Christie & Alkin, 2013; Greene, 2016; Hood, Hopson, & Kirkhart, 2015; Hopson, 2009; House & Howe, 2000).

Limitations

The foci of these three papers are aspects of ASER's design—the goals that underlie its design, the politics and needs that informed it, the methodological decisions that define it, and the influence its design has had globally. In these papers, I relied heavily on analysis of the

ASER architects' interviews, observations of the ASER process, and analysis of ASER documents, especially training protocols, ASER reports, and blog posts written by ASER participants as they underwent training and collected data. Data from interviews with ASER volunteers, partners, and state-level staff members will be utilized in future manuscripts.

Research Significance

The goals of this research on ASER are two-fold. The first goal is to better understand how evaluations matter to the improvement of social systems (like education) in India, specifically, and non-Western democratic countries, more generally. Within the field of evaluation, there are debates over how to conceptualize and study the impact of evaluations on societies. Furthermore, the literature on *evaluation influence* (Kirkhart, 2000) lacks empirical studies that explore and test the current frameworks that have been developed to conceptualize the relationship between evaluation, democracy, and "social betterment" (Henry & Mark, 2003). My research, a case study for examining the influence of a large-scale evaluation of education in India, addresses this need. Specifically, this dissertation examines ASER's design (i.e., its theories, approaches, methods, process, and findings) as a source of its influence.

The second goal relates to the limitations of evaluation theory. Theories of evaluation tackle issues of values, cultural responsiveness, and inclusion in evaluations, but they have been theorized from Western contexts. At the same time, evaluation scholars emphasize the importance of conducting evaluations "from a culturally competent standpoint" (Chouinard & Cousins, 2009, p. 458) that reflects knowledge of a context's power dynamics and social inequities. In India, there is currently minimal documentation on how evaluations are conducted (Hay, 2010) and how they are designed to address issues of culture and equity. Moreover, though evaluation theory frequently addresses the evaluation of smaller-scale programs, in reality,

evaluations increasingly tackle giant social systems and are charged with producing system-wide data for public consumption—data that should be credible and valid across culturally diverse, sub-national contexts. My research engages with evaluation theory in India amidst growing interest in evaluating its social systems for accountability, transparency and evidence-based decision-making.

PAPER 1. "We Went in Favor of Rebellion": A Narrative Inquiry into the Design of India's Annual Status of Education Report

Abstract: The *Annual Status of Education Report* (ASER) investigates primary education across rural India via surveys and assessments. ASER is groundbreaking in its scale, cost effectiveness, simple design, and "citizen-led" approach. The non-governmental organization, Pratham, created ASER in response to children's learning issues and as a "proof of concept" to Indian policymakers. ASER's origins in grassroots programming and policymaking distinctly shape its methodology and goals. Describing its development amid large policy shifts in India's school system, the narratives of ASER's three main architects explain how ASER forged a path complementary to government efforts while promoting citizens' rights to independently evaluate learning.

We
People of India
From different states and regions
Speaking different languages
Sat with our children
And looked
Within
Inside our homes
At our villages
Into our schools
And prepared this report
For ourselves
To build a better India
—ASER 2005

The *Annual Status of Education Report* (ASER), which investigates the condition of primary education across rural India, is produced via a massive survey and assessment process. In the sphere of educational monitoring and evaluation (M&E), ASER has garnered international attention because of its groundbreaking scale, cost effectiveness, simple design, and "citizen-led" approach: it engages over 25,000 volunteers in conducting ASER each year (Banerji, 2013, p. 5). Over a decade ago, ASER began in two ways that coalesced into one big idea: the first was as an extension of the basic education programs ran by Pratham, an Indian non-governmental organization (NGO) focused on children's learning. The second was as a giant "proof of

concept" pitched by Pratham leadership to Government of India (GOI) officials in order to generate data on India's education system. Thus, according to Dr. Rukmini Banerji (personal interview, May 20, 2015), one of the architects of ASER, ASER materialized from the intersection of "a ground-level view" of education with "a bigger policy, macro angle." Dual origins in grassroots programming and national policymaking shaped ASER's methodological features and goals.

This article offers a history of ASER's development as told through the narratives of its three main architects, Dr. Rukmini Banerji, Dr. Madhav Chavan, and Dr. Wilima Wadhwa. Following the introduction, the second section outlines this study's questions and methodology. In section three, Dr. Banerji narrates ASER's emergence from Pratham education programs. In section four, Dr. Chavan describes how ASER formed through engagement in policymaking. Dr. Wadhwa explains the methodological choices that comprise ASER's design in section five. The conclusion captures how the architects' interweaving accounts tell a spirited story of ASER's creation amid the substantive policy shifts and educational changes underway in India. Finally, the postscript highlights the secondary argument of this paper: the power of ethnography and narrative inquiry as methodological techniques for studying the history behind and values underlying monitoring and evaluation efforts. Because when one knows the story of something, when one can trace its origins from conceptualization to experimentation to realization, one can see it more fully and interpret its meaning anew.

Study Questions and Methodology

A 10-month ethnographic study of the entire ASER process was conducted with fieldwork in Delhi and three states: Rajasthan, Manipur, and Tamil Nadu (see Appendix A for more information). Drawing on the larger study, this article addresses the following questions:

1. How do ASER's origins in policymaking and grassroots programming shape its methodology and goals?
2. What are some implications of the design choices made for ASER?
3. How do these choices influence ASER's importance as a "proof of concept"?

These questions are answered primarily from semi-structured interviews with Drs. Banerji, Chavan, and Wadhwa, who all consented to be identified in this article.[5] Analysis of their interviews reflected a narrative inquiry approach that underscores the importance of social location, life experience, and individual nuance in knowledge creation (Rhodes & Brown, 2005). Data gleaned from the ASER architects' interviews are intrinsically tied to their unique positions and experiences, so their collective storytelling constitutes one among many stories that could be told. In addition to member-checking quotations and interpretations, ASER reports, documents, and publicity materials were utilized to supplement and triangulate their accounts.

This study of ASER is framed as research on evaluation, which is defined as "empirical research on evaluation conditions, processes, and consequences" (Cousins et al., 2015, p. 80). My research on evaluation specifically employed a narrative inquiry approach to understand the "conditions" that prompted ASER's creation and the decision-making "processes" that undergird its design. Nevertheless, why does research on evaluation matter? Christie (2012) identifies four key areas where research on evaluation can make important contributions, one of which is supporting "the development of knowledge on evaluation issues" (p. 5). This study highlights two important evaluation issues: 1) methodological choices: the rationale and decision-making processes behind ASER's methodological features and 2) evaluation goals: the vision and priorities that shaped ASER's innovative design. Cousins et al. (2015) note the importance of

[5] Drs. Banerji, Chavan, and Wadhwa were individually interviewed in Delhi for one to one and a half hours each in May 2015.

"reflective case narratives" in research on evaluation, and they highlight the significance of narrative inquiry as a methodology for constructing knowledge about evaluation: "Narrative inquiry is a rich and valued approach to understanding the complexities of [evaluation] practice through the eyes of those immersed in its context" (p. 80). "Through the eyes of" ASER's architects, "the complexities" of their design practices are revealed, offering unparalleled insight into the dilemmas they faced in trying to achieve their goals for the evaluation, while bounded by the limitations and considerations of their context: the expansiveness, diversity, and (often) remoteness of rural India. Despite the nation's size and global importance, research studies on evaluation in India are rare. Thus, this paper and larger study address a substantial gap by profiling a pioneering evaluation model designed by leaders in Indian civil society.

Origins in Pratham Programming

Much has been written on the persistent challenges of UPE in India (see as examples Drèze & Kingdon, 2001; Kingdon, 2007; PROBE Team, 1999). At independence, India had a 39% nationwide literacy rate and "began transforming the elite education system inherited from its colonial past into a mass education program" (Govinda, 2008, p. 431). While important strides have been made, there remain economic, social (e.g., gender, caste, and religious), and regional (e.g., remoteness and infrastructural) reasons why children do not attend school and complete a primary education today (see Govinda & Bandyopadhyay, 2010). Moreover, immense linguistic diversity complicates providing equitable instruction and accessible learning materials to all Indian children.

In 1995, Pratham was established as an educational NGO in Mumbai (i.e., Bombay) with a focus on providing "high-quality, low-cost, and replicable interventions to address gaps in the education system" (Pratham Education Foundation, 2016). As one of India's largest NGOs,

Pratham's mission is to "improve the quality of education in India and ensure that all children not only attend but also thrive in school" (Pratham Education Foundation, 2016). Pratham programs that aim to close quality and equity gaps in basic learning for primary-age children in underserved rural and urban communities directly aid in India's universalization of primary education (UPE). Rather than circumventing government efforts, Pratham tries to "supplement" primary schools at a large-scale to "reach as many children as possible" and to "create an adoptable demonstration model for state governments" (Pratham Education Foundation, 2016). Pratham's strategy is to strengthen states' capacities to provide good public education by developing novel school-based learning interventions, which illustrate alternative instructional approaches and are feasible even within the constraints of the government school system.

Dr. Rukmini Banerji, Pratham Chief Executive Officer (CEO), has been involved in the NGO's work for two decades. Recalling Pratham's early days, she describes the idealism in how people thought about schooling:

> Universal enrollment was a big goal—it was a high priority at the macro-level as well as all the way down to [the] community or village. Parents, teachers, schools—*everyone* was kind of aware of this as a big issue, and the broad understanding that our children should go to school was very widespread in India even 15 years ago. (R. Banerji, personal interview, May 20, 2015)

Meanwhile, the government's focus was on expanding school infrastructure and training more teachers to absorb new children into India's education system. Dr. Banerji (R. Banerji, personal interview, May 20, 2015) describes Pratham's efforts "in Bombay in the late 90s" where it was "working together with the city" government:

> [We] were already pinpointing areas where this [school enrollment] was a big problem…. There was also a specific focus on types of children who were likely to be out of school so…the problem was well understood and accepted, and a whole variety of solutions for how to achieve very high enrollment were in place. That is not to say that the last mile had been conquered, but at least you had a sense of where it was, who they were, what the difficulties were, and some attempts at solving the problem.

Pratham supported the local government's efforts to enroll more children by identifying city regions where children were under-enrolled in schools and locating what kinds of children were not enrolling—presumably, children of certain gender, caste, tribal and religious identities or with difficult economic and social circumstances (e.g., child labor and child marriage). Dr. Banerji (personal interview, May 20, 2015) concedes other problems beyond enrollment were less noticeable:

> The enrollment—every child in school—is a visible problem, but every child not learning is an invisible problem. The assumption in many parts of the world including India is that school will lead to learning: I mean it may not lead to as much learning as you like, but it's going to as soon as you have everyone in school and teachers…. [Pratham workers] were beginning to see that this does not automatically lead to that: You could be in school, and you could be going everyday but you are not learning as much—what is not as much? What is much? It is a little fuzzy. You have a feeling that things are not going that well, but there is not a real picture of what that means.

India was striving to realize *Education for All*, but Dr. Banerji expresses that people everywhere were operating under a false truth—that schooling equaled learning. Through its programs, Pratham workers were encountering children who did not know what one might expect given their grade-level and school attendance. This early observation was "fuzzy," but Pratham staff detected the possibility of children enrolled in school, but "things…not going that well" with their learning.

Confirming Learning Problems

In 1997, given their mounting concerns about learning, Pratham begins testing children as Dr. Banerji (personal interview, May 20, 2015) explains:

> We did a little assessment of math in 20 schools in a part of Bombay that was not particularly badly off. It wasn't like a famous slum…what we found is that children in 3rd and 4th grade were having difficulty with the basics…. We put this all in one kind of sheet, and I remember coming back to the office one late evening and saying, you know, "this needs a math revolution."

Assessing children led to a disturbing discovery: many government-school students struggled to do basic math like recognize numbers and do double-digit subtraction.

Sharing "Easy-to-digest" Data

Pratham's supportive disposition towards the government enabled the NGO to form a partnership with the city in pursuit of this "math revolution":

> In a month, the city government had launched a big math program, so we also saw that very easy-to-digest, basic data and a receptive government could actually do a lot of things. We had some experience of [success], and they called that program…"Jump to 100"… What we looked at was, do kids know their numbers up to 100? Can they do their addition and subtraction? [We had] the experience that a problem…when put into a concrete form that is easily digestible by many people, can lead to solutions. (R. Banerji, personal interview, May 20, 2015)

Two factors in Pratham's success were "very easy-to-digest, basic data" and "a receptive government" that embraced Pratham's collaborative approach toward its UPE efforts. Because Pratham tests measured arithmetic skills that education non-experts could readily identify as foundational to children's mathematical learning, Dr. Banerji just straightforwardly reported the number of children that could perform each skill out of the total children tested in 3rd and 4th grades.

Using Data in Educational Interventions

The math assessment dispelled assumptions and led to new assessment practices within Pratham:

> We [determined]…there were some assumptions…that were perhaps getting in the way of recognizing a new problem…that the children were not learning, so anywhere where we worked—whether it was urban or rural—we would first do a kind of census of the kids we wanted to work with whether it was in one school or in one village. (R. Banerji, personal interview, May 20, 2015)

Exclusively focusing on enrollment obscured whether children were actually learning in school. Thus, Pratham tested all children in basic reading and math skills wherever they planned

programs. This "census" created baseline data from which Pratham workers could bulid upon children's current skills.

Employing data to guide its learning interventions, Pratham devised an approach called Combined Activities for Maximized Learning (CAMaL) that quickly developed children's basic math and reading skills:

> By 2002-2003, we had early versions of what we call CAMaL, which said, "If you do these five [learning] activities every day for X number of days, most kids will reach reading fluency of this type if it is in a language that they are really used to already, and if the kids are eight and above." We had a solution for at least basic reading and basic math. (R. Banerji, personal interview, May 20, 2015)

Assessments became vital to Pratham's work because "[f]undamental to the CAMaL approach is the grouping of children on the basis of their learning levels rather than on the class in which they are enrolled."[6] Hence, CAMaL was introducing a type of data-driven, student-centered approach to instruction wherein all children, regardless of their grade-level and age, were being taught through small-group activities that built upon their current competency-levels.

Engaging Communities in Learning

By 2003-2004, Pratham was testing children, compiling test data, and sharing it with rural communities in "village report cards." Dr. Banerji describes their discovery that disseminating test data was powerful. Proof and clear articulation of children's learning problems proved vital to Pratham gaining local support for its interventions:

> We were working in a lot of rural areas doing these village report cards, and…we also had a solution that we could offer so it wasn't that we were just highlighting a problem and moving away. Wherever we worked, we felt that the problem articulation was important because then you got engagement…because it is the people who engage who will be partners in the solution. (R. Banerji, personal interview, May 20, 2015)

[6] For more information on the CAMaL methodology, see Pratham Education Foundation (undated).

Proof of learning problems came not only from sharing test data with villagers and the village head (i.e., Pradhan or Sarpanch), but also from people witnessing children being tested. In a 2013 article entitled "The Birth of ASER" published in *Learning Curve*, Dr. Banerji describes going household to household to ask if children are home, to see if they go to school, and to test them in reading and math. Watching testing and discovering what their own children could or could not do was persuasive. Dr. Banerji (2013) describes one villager's reaction:

> I remember an old lady shaking her head and saying "this is not a survey", "Why not?" I asked her…. She said, "a survey is when you don't know but we know. A survey is when you come from the city to find out what we village folk are all about. But this is not a survey. Because you don't know, and we don't know and in fact, even the children don't know what they can do. It is only when you ask them to read or do sums that we all find out. We find out together and we realize what our reality ("asliyaat") is." (p. 86)

Dr. Banerji asserts that when people had the opportunity to directly participate in a learning solution, the problem articulation itself became more meaningful. Days after assessing, Pratham workers returned to the village to share an official "village report card" and facilitate dialogue on what can be done based on the findings (Banerji, 2013). Dr. Banerji (personal interview, May 20, 2015) explains:

> We had methods and mechanisms that we knew worked to suggest: "here is the solution." This is what we did at the village level. Some of these Pratham programs had been evaluated by this time by J-PAL [The Abdul Latif Jameel Poverty Action Lab], so we also knew that independent evaluations were showing that some of the solutions for improving basic learning were effective—not just our own measurement.

Pratham's confidence in its village report cards and learning programs grew as outside research organizations studied Pratham interventions and provided evidence of their effectiveness (see Banerjee, Cole, Duflo, & Linden, 2007).

Making Learning "Visible"

Dr. Banerji describes ASER's "backdrop" as rooted in Pratham's view that its interventions were only one strategy for addressing India's learning problem. Pratham believed

people could take a variety of actions based on its evidence; its main goal was to raise awareness and prompt public dialogue by making the learning problem "visible":

> We felt that the problem was not visible enough, which is why people were not looking at it and if it became visible then action on it [would come]. It would be helpful at [this] stage not just to say, "Hey, here's a problem, guys: go and look at it," but to say, "Okay, we have one kind of solution—you may not like the solution, come up with your own." That is something that we feel works…. The ASER reading tool came out of this process. It lends itself very well to the action that we follow; it doesn't have to be the action that you follow, but it's not an assessment or an evaluation tool alone—it is step one in an instructional process. You may not use the process, but for us it was a segue from one to the other. (R. Banerji, personal interview, May 20, 2015)

ASER highlights basic learning issues in a way that Pratham interventions logically follow as a solution: Dr. Banerji acknowledges that ASER's testing tool is really "step one" and "is not an assessment or an evaluation tool alone." However, ASER did evolve into a stand-alone inquiry process that produces data interpreted apart from Pratham solutions.

Dr. Banerji's account outlines ASER's grassroots origins in Pratham programming, but ASER's beginnings can also be tied to macro-level engagements. Dr. Madhav Chavan (personal interview, May 20, 2015), Pratham founder and former CEO, explains ASER's origins in a funding proposal he submitted to an international NGO in the Netherlands:

> I think this proposal was being written in 2003 as to what we wanted to do… [Pratham] basically said that we could summarize and do a quick survey of the areas that we were working in, and we proposed [it] to the governments in several states of India…I remember writing we could submit an "annual status of education report"…so that thought was there—that this was likely to go pan-India, but we had not quite conceptualized how it came out later on.

This proposal articulated a new idea wherein Pratham would build upon the village report card concept. Dr. Chavan proposed expanding Pratham testing from the village-level to a higher level and to garner support from state governments for creating this pan-India baseline of learning data. He mentions calling it an "annual status of education report," but the idea was still underdeveloped.

Origins in National Policymaking

Proposing ASER to the Indian Government

Dr. Chavan (personal interview, May 20, 2015) identifies a change in India's central leadership as significant to ASER's development because it led to shifting national priorities and an advisory role that enabled him to propose ASER to people with power.

> The big change happened when in 2004 a new government took over. Dr. Manmohan Singh and his finance minister P. Chidambaram started announcing that they were in favor of outcomes and not just outlays [e.g., expenditures and infrastructural inputs] to see the effectiveness of their government. I happened to be in…the National Advisory Council appointed by the Prime Minister. Sonia Gandhi was the chairman of that committee.

Dr. Chavan reveals he had unique access to Prime Minister Manmohan Singh and the head of India's Congress Party, Sonia Gandhi. He emphasizes that Prime Minister Singh was already thinking about a focus on outcomes in monitoring government activities, creating an opening for Dr. Chavan to pitch ASER as an important undertaking. Speaking to the Prime Minister and the Deputy Chairman of the Planning Commission Mr. Montek Singh Ahluwhalia, Dr. Chavan (personal interview, May 20, 2015) recommended that the central government share more information about the education system with the public:

> I suggested that the government should do an "annual status of education report" because the government had levied what was known as "the 2% cess" [i.e., levy or tax] for the MHRD [Ministry of Human Resource Development]….Everyone was paying towards it.

MHRD recently began receiving funds from a 2% income tax that was earmarked specifically for supporting India's primary education system. With the GOI's new economic reserves in mind, ASER began as a proposal for the *government* to undertake this data collection effort or, at least, to directly support ASER in a large way.

Despite Dr. Chavan's proposal, the responsibility for who would execute and fund an "annual status of education report" quickly shifted back toward Pratham:

> The feeling there was that the government was in favor of the *idea*…but Dr. Manmohan Singh and Mr. Montek Singh Ahluwhalia… feared that if *they* entered the picture by way of funding or otherwise, then we would have to carry a huge baggage of the government bureaucracy: their experts who would not want to take this kind of risk in the program and all of that. He said, "Why don't you go ahead and do it. Show us a proof of concept; show us how you would do it." And, we said "okay." In 2004, even before the ASER survey, we did a quick assessment of how we would go about doing it and came back with a quick report of about 120 districts. That was the beginning. It was not called anything; it was not published, but we did it as a proof of concept and we presented it to the Planning Commission people, saying that this can be done regularly (M. Chavan, personal interview, May 20, 2015).

Dr. Chavan describes the 2004 pilot process of ASER as an experiment (a "proof of concept") for GOI's planning and monitoring body, the Planning Commission, which assesses the nation's resources and formulates five-year plans for "the most effective and balanced utilization" of resources and accordingly, monitors progress in order to recommend "adjustments of policy" (see Government of India, 2014b). Pratham did the pilot to prove that the government could feasibly collect annual, nation-wide data on children's learning to inform policy decisions. Post-pilot, Dr. Chavan explains that the Commission was not keen on getting involved:

> This is when they say… "This looks okay and good, but if the government enters it, then the innovative part of this will just die, so then why don't you just do it? I mean, you say that it won't cost a whole lot of money…. if you get the government involved, then you will be carrying a lot of baggage." And so, we said "Okay, then why don't we just do it ourselves?" …. There was a whole thing about should we or should we not (M. Chavan, personal interview, May 20, 2015)?

Given the possibility that government involvement would impede ASER's execution and damage its merit, Pratham leadership was tempted to undertake ASER on its own but anticipated some risks in doing so:

> There were some concerns that if we did it without the government, and the report came out negative, then how would the government respond to it? But, we went in favor of rebellion, and said you know, "why not? We have a right to assess our children and see what it means." And so, we thought of doing this report…it had to be called something, so the name *Annual Status of Education Report*, the acronym [ASER] fit very well with the Indian word "aser" which means "impact," in fact. We didn't know it was going to be annual; we didn't even know if we were going to be able to do the first one, but we did it. (M. Chavan, personal interview, May 20, 2015)

Through this series of interactions ASER was finally born. ASER's end product would be a report of children's basic learning levels in reading and arithmetic, and its process of data collection would be distinctly independent of the government both operationally and financially. "We went in favor of rebellion" suggests that Pratham recognized its right as a civil society organization within a democracy to independently evaluate the efficacy of public institutions like schools. Emphasizing that India's government is not a monolith, Dr. Chavan (M. Chavan, personal interview, May 20, 2015) explained ASER did not have widespread support:

> It was the support of one or two people in the Planning Commission…[they] said that "we are with you." In fact, a couple of Planning Commission members also gave *money*, which we were collecting [to fund ASER]—and they said, "this is our voluntary donation for it." So, they were quite enthused about it.

The exchange of ideas instrumental in shaping ASER involved only a few GOI officials within one department. Why did Dr. Chavan pitch ASER to the Planning Commission instead of the MHRD or National Council of Education, Research and Training (NCERT) that directly oversees India's school system? Perhaps, he took the opportunity to propose ASER to people with whom he had influence and decided that the Commission's economists and planners would be a sympathetic audience to ASER's large-scale quantitative data, which could be used for crafting national priorities and monitoring (Government of India, 2014b). When compared with NCERT and MHRD, Pratham leadership possibly viewed the Planning Commission's use of educational data as more influential on India's school system.

Addressing Data and Accountability Needs

Dr. Chavan (personal interview, May 20, 2015) describes two issues that defined India's educational context that illuminate how a need for ASER was identified: 1) the recent changes in national school policy and 2) the educational data available for decision-making.

> The Government of India had set up a process [for collecting data]—this is the DISE, District Information System for Education, but they were not doing it too well. I mean their reports were two years old; information was not coming from all the states, so there was no energy in it. I think that the DISE people themselves were complaining that the states don't cooperate. Also, I don't know if [you] remember that Sarva Shiksha Abhiyan [SSA], this Education for All [EFA] program that the government had set up in 2001-2002, did not have too much money. But now, with the new education cess, SSA was going to get a lot of money, so there was a system, but it was very weak from many points of view. Good data, enough data, correct data, timely data—I think they were very weak in that.

Dr. Chavan identifies a disconnection between policymaking and using "good" data in the years preceding ASER. The GOI apparently recognized data's importance in making informed policy decisions and in monitoring policies' implementation because they initiated DISE and produced reports of DISE data (though it was fraught with issues). DISE is a main source of education system data collected by India's district governments in every state. The National University of Education Planning and Administration (NUEPA), which compiles DISE data that are self-reported by states, warns it is not involved in verifying the data's "truthfulness" or "accuracy" (see disclaimer in National University of Educational Planning and Administration, 2013). Several issues with DISE data persist (see Govinda & Bandyopadhyay, 2011). One impetus for starting ASER was the realization that GOI educational data was poor and impeded SSA's success. (SSA is a set of initiatives to facilitate the expansion of primary schooling in India and encourage all children to enroll and attend.)

In context of EFA, significant policy and funding shifts were underway in India. Efforts emphasized universal education and gender parity by addressing gaps in educational infrastructure and access. In 2004, SSA was the GOI's clearest effort toward universalization, but according to Dr. Chavan, it was initially underfunded. With the 2% cess, Dr. Chavan's statement suggests the system's weakness changed from underfunding to accountability: how well was SSA implemented and the new tax money spent?

Multilateral organizations like the World Bank were also funding universalization efforts, which Dr. Chavan (personal interview, May 20, 2015) believed adversely affected data quality and led to problematic approaches to assessment:

> The second [issue] is that [the Indian national government] had, I think as part of the World Bank money that came in or otherwise, there was a system of doing purely assessments of children. These were based on pen-and-paper tests, more provisional than anything, and, they did not test reading. What they surmised was that if [children] could answer the questions correctly, then obviously they could read, and it was all multiple choice…. That too was not very well done…. No one was taking education data seriously. It was just not available.

No one was "taking education data seriously:" on the one hand, test data were not made readily available, and on the other hand, the test data were anyhow invalid because they did not establish whether children could actually read the multiple-choice questions and thus, did not insure that test scores were reflecting children's competencies on the intended content. As a result, Dr. Chavan believed that Pratham could do a better job in producing meaningful data on the educational system's quality and on children's learning: "Government was [not] collecting something in a way we can do in that [educational data] was not something in the public consciousness, the government itself was not looking at the data carefully" (personal interview, May 20, 2015).

Pratham leadership surmised that if they were independently collecting learning and education system data through ASER and designing the tests and disseminating the data themselves, they could share data openly in ways that ordinary citizens could easily interpret in context of the India's goals under SSA. If government's data were not sound or taken seriously, ASER could fill the gap and create public pressure for the GOI to evaluate the education system through evidence. Therefore, ASER had accountability aims of openly sharing credible data and initiating transparent conversations about government's provision of primary education.

Development of ASER's Methodology

Forming a Design Team

Once Pratham leadership committed to ASER, there was a small window for designing its methodological approach. In addition to Drs. Banerji and Chavan, the main architect of ASER's methodology was Dr. Wilima Wadhwa, a statistician with expertise in development and econometrics: "A lot of my research has to do with working with large-scale data because I work on India" (W. Wadhwa, personal interview, May 15, 2015). As an occasional "academic resource" on "Pratham research committees" (W. Wadhwa, personal interview, May 15, 2015), Dr. Wadhwa eventually contributed a great deal of the technical knowledge needed for designing ASER. She explains:

> So 2005, I remember [in] October, Madhav said, "how would you like to design a nationally representative survey of learning?" I was like, "oh, my god, that would be the most exciting thing to do!" Because up until now, I was a user of survey data, but I am not looking at it from the other side—how would you actually design [something] if you wanted to answer a question? How would you actually design a national assessment, which was representative at certain levels? So, that is how I got sucked into it. He made me an offer that I could not refuse. (W. Wadhwa, personal interview, May 15, 2015)

Dr. Wadhwa got "sucked in" by the intellectual challenge of designing ASER but also believed Pratham was well positioned to undertake it with "a fair amount of work [done] on learning outcomes" (W. Wadhwa, personal interview, May 15, 2015). Her account also emphasizes there was little time to devise ASER's methodology and process:

> We were having this conversation in October…. So, we [still] had to design it. We had to get the teams together. We had to actually *go*—I mean…we had *nothing*. We were having a *conversation,* okay? And, ASER was published…in January of 2006. (W. Wadhwa, personal interview, May 15, 2015)

ASER 2005 would be publicly released with less than four full months from methodological design to publication. In the coming weeks, the ASER design team—Drs. Banerji, Chavan and

Wadhwa—made several methodological decisions (covered in the following sections) that established ASER's innovativeness.

Defining and Measuring Learning

Dr. Wadhwa explains that the first decision was determining how to define and measure learning: "What is it we wanted to measure? We wanted to measure *learning*. Well, *what* is learning? How do you…even define learning?" (W. Wadhwa, personal interview, May 15, 2015). Experience with Pratham assessments directed how they chose to measure learning, thereby defining learning in ASER as basic skills in reading and math:

> We were fortunate because the ASER [test] tool that we use was something that existed, and Pratham had been using it at large-scale, and they had been actually using it as a *teaching device*…to figure out where children were at…The dilemma was we could have designed an assessment which was grade-specific…[and] in different subjects, but we said, "We are doing it for the first time. Maybe it turns out that everybody could do this"…. [But] given that we were going to generate the first estimates of learning, [we said,] "Let's keep it simple." (W. Wadhwa, personal interview, May 15, 2015)

Pratham's testing tools are one-page reading and one-page math assessments that are designed to identify each child's level of basic reading or math development (see Figures 1-2). For example, if a child can identify letters on the reading test, but not read simple words, that child is at "letter-level" reading proficiency. If a child can do double-digit subtraction successfully on the math test, but not complete the long division problems, that child is at "subtraction-level" math proficiency.

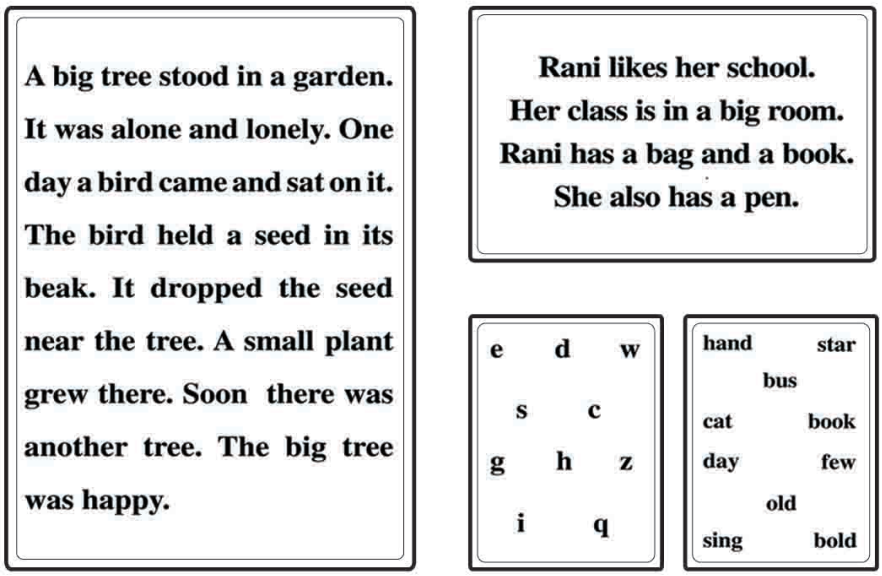

Figure 1. Reading Test (English). From "ASER Reading Tools," by ASER Centre, 2017, retrieved from http://img.asercentre.org/docs/Aser%20survey/Tools_Testing/Reading/english.jpg. Copyright 2017 by ASER Centre. Reprinted with permission.

Figure 2. Arithmetic Test (Hindi). From "ASER Math Tools," by ASER Centre, 2017, retrieved from http://img.asercentre.org/docs/Aser%20survey/Tools_Testing/Maths/hindi.jpg. Copyright 2017 by ASER Centre. Reprinted with permission.

The architects suspected that many rural children across India did not possess the basic skills required for higher-order learning and doing grade-specific tasks, so they decided to use Pratham assessments to create a baseline of learning data and offer a pan-India snapshot of what children know how to do. Testing very basic skills across such a large age-range (ages 5-16 years old)

was a calculated risk based on Pratham's hunch that even older children were struggling with basic reading and math.

There were many advantages to adopting Pratham's tests for ASER including the tests' apparent validity over time and across contexts. During the village report cards, the testing tools were useful for having discussions with parents about their children's learning. According to Dr. Wadhwa (personal interview, May 15, 2015):

> Given the simplicity of the tool, if we show [it] even [to] a mother who has not been to school, she can look at those four lines [of the paragraph] and *figure out* that what is the level that we are talking about, and understand that "oh my god, my child cannot read *this*." [It is] simple for parents to understand and not just parents, but parents who have not been to school. Because we know that 50% of rural children have mothers who have never been to school [and that] 20% of rural children are first-generation learners—*both* parents have not been [to school].

The tools seemed to possess real explanatory power even for illiterate parents. Dr. Wadhwa expresses the tests had face validity—both credibility and meaning—for rural families.

Keeping it Simple

The architects were motivated by a key principle in designing ASER—"Let's keep it simple"—so ASER could be easily interpreted by multiple stakeholders. Dr. Wadhwa (personal interview, May 15, 2015) explains that they wanted ASER to be "simple to disseminate:"

> You are not talking about some complicated test [so a] person in the government who is not necessarily in the high echelons of policymaking, but is actually doing the work, can understand the results of that survey. So, it wasn't just that we were trying to get that guy on the top who might be able to understand complicated results, but you are talking about *all* levels of the government.

The simplicity of ASER assessments facilitated sharing the results with all people from the village head up to the block-level, district-level, and state-level officials. ASER tests must produce results easily aggregated and interpreted via report tables and graphs. This design principle also facilitated widespread public interest in ASER:

> *Simple* because it would inform various [things]: One, if it is simple, it is easy to do and given that one of the main things of the ASER model is this whole *citizen-led* thing: it can be administered by anybody [and] simple, so it could be *understood* by everybody. We did *not* want another survey, which was another academic survey to be used by *experts*. It should be when you disseminate [the findings], *everybody* should be able to *understand.* (W. Wadhwa, personal interview, May 15, 2015)

Keeping ASER simple would enable ASER findings to be shared by mainstream national and state media and used by academics, Indian NGOs, and international aid agencies. The simplicity of ASER also helped achieve a defining aspect of ASER's design: its "citizen-led" approach, which Pratham leadership defined as ordinary citizens, not educational experts, empowered to use ASER tools accurately to test children.

Choosing Testing Site and Format

The ASER architects considered where testing should be conducted and what format the tests should have. Dr. Wadhwa (personal interview, May 15, 2015) explains that typical pen-and-paper assessments in schools systematically omitted some children:

> Typically, [when] learning assessments are done, they are pen-and-paper assessments done in school, even the Indian government's assessment…. Now, that was I think one of the most *crucial* decisions that we made at that point. If you test in *schools*, then you are limited to a particular kind of school…. See, the reality in India is that children go to many different kinds: they go to public schools; they go to private schools; they go to religious schools, so how could we capture all these kinds of schools, apart from the technical difficulty of having a *listing* of these schools? And, *most* importantly is that while enrollment levels were *high,* close to 95% even in 2005, we *know* that children don't attend school regularly. So, if you showed up in schools, even if you could do that perfectly and get permissions and all of that to enter schools, if you went one random day, you would miss out on a *huge* number of kids. In UP [Uttar Pradesh] and Bihar, 50% children are not attending school regularly.

There is no official list of some types of schools like unrecognized schools and madrasas, so they cannot be sampled. Gaining entry into and cooperation from these schools in order to test children was an issue and so was children's irregular attendance. Dr. Wadhwa identified these issues of representation as major threats to the accuracy of data that a school-based assessment

would produce: “You don't want to have a special assessment of children who only go to government schools, or children who only go to private schools, or children who [are] only attending regularly. Because you are going to get a biased estimate” (W. Wadhwa, personal interview, May 15, 2015). The architects knew that administering ASER tests in schools would skew the data, only reflecting certain children’s learning and not the children *mos*t in need:

> *That* creates this whole self-selection problem because if you want to bring this whole learning crisis to the fore and *do* something about it, who are your *target kids*? The kids who are actually *behind* and these are likely to be kids who are not regularly attending. So, you would miss out on them completely in a school survey. So, the idea was we want to capture *all* children. (W. Wadhwa, personal interview, May 15, 2015)

Thus, the architects avoided schools as the main mode for sampling children and decided to test children in their households, so households became the mechanism for sampling children for ASER testing. Dr. Wadhwa expresses this was a “a very, very key decision.” She reveals, “All of this was…tied to the fact, that…in some sense, we were the forerunners here. Nobody had ever done this, and if you want to do this, you want to capture *everybody*.”

To accurately capture every child’s learning ability, the architects also had to modify how tests were normally administered:

> [We were] thinking “outside the box” [rather than] this traditional thinking of pen-and-paper testing in school. That brings us to, how would we administer the test? How do you give a pen-and-paper test to a child that cannot read? So, it had to be a one-on-one testing. (W. Wadhwa, personal interview, May 15, 2015)

ASER one-on-one testing would be done orally with the ASER surveyor sitting beside a child, giving verbally instructions on how to complete each part of the test and observing the child’s ability to fulfill each task. This innovation increased the math and reading tests’ validity because the surveyor could be confident that if the child was having difficulty completing the task, it was not because she or he could not read the directions or had undetected mother-tongue barriers.

Sampling for District-Level Representativeness

Dr. Wadhwa (personal interview, May 15, 2015) explains that the ASER sampling decisions were motivated by "at what level" they wanted to have influence on educational decision-making:

> *There* we decided…that we wanted estimates at the district level. Now, *why* at the district level? Because *educational plans* are made at the district level. And again, this was coming from the fact that we wanted to do something. This was not data that we were generating for researchers or experts. You want to *do* something; you want to disseminate it, and you want these findings to feed back into education plans…. So, we said, "okay, that is the lowest level at which we wanted estimates." That decided really the design of ASER: [it] is really at the district-level and you just blow it out—in terms of how many districts you cover. The sample size and all of those decisions are made at the level that you want representation.

Dr. Wadhwa emphasizes that the primary goal for ASER data was its use in educational planning, not academic studies, which made ASER's priority of developing a sampling frame that would ensure data representativeness at the district level relatively straightforward. Alternatively, if producing data for academics had been ASER's focus, the architects may have felt torn about whether researchers would want results at the state or regional level, or whether they would seek data representativeness at much lower levels of governance, such as district sub-levels like the block, cluster, or village. Moreover, if the architects had been most interested in ASER data use to design community-level or school-level interventions (like Pratham programs), the ASER sampling frame would have to be adjusted toward representativeness more locally at ideally the village level, which would increase ASER's data collection burden significantly by requiring a much higher number of households sampled and children tested (at each age). Dr. Wadhwa determined that sampling 20 households in each village for 20 villages in each district (totaling 400 households per district) would be sufficient to achieve district-level representativeness. In this sampling design, the number of children tested would vary by village (and thus, district) because children (5-16 years old) would be tested only if they lived in a

randomly sampled household: systematically sampling households to reach the 20-household threshold would typically yield some sampled houses that are childless and others that contain many school-age children.[7] Finally, following ASER 2005, Dr. Wadhwa decided that districts would have 30 randomly sampled villages that rotated: 10 new villages would cycle into the sample and 20 villages would carry forward from the past year's sample (10 villages that were sampled from the past year plus 10 villages sampled from two years ago), which she surmised would enable Pratham to use ASER data to analyze change over time while still maintaining necessary breadth in the diversity of villages being represented.[8]

Determining Frequency and Timing

Next, the architects recognized they needed to determine the frequency of ASER and the timing of its data collection, which would have further implications for ASER data's use in educational planning. Dr. Wadhwa (personal interview, May 15, 2015) explains why frequency and timing were crucial strategic decisions:

> Then, it was *frequency*. Okay, how often did we want it? And, when during the school [year] should it be done? And, there again…I mean *annual* is part of our name—ASER is the *Annual Status of Education Report*—so…we *went* in with the commitment that we were going to do it for five years. It was *built* into [the ASER design that]…it is not a one-off survey. You are going to *track* it, year on year on year…. So given [that]…the idea was…we didn't want to [just] say, "children can't do this," [but] we wanted to figure out, what is it that children *can* do? …. It is very easy to say, "Oh, we went in with Standard II [reading tests] and nobody could do it." Then, *what*? That doesn't tell you *anything* right? …. This has to generate information, which will *feed* into not just policy and decision-making, but feed into *strategies* of how to improve the thing. You have to know what is it that children can *do*, *where* they are to then *take* them to the next level.

[7] The ASER 2005 sampling design and key sampling decisions, like employing the GOI 2001 Census lists to randomly select villages "within each district using probability proportional to size method of sampling (PPS)," are described at length in the report (Pratham Education Foundation, 2006, pp. 5, 127-128).

[8] See the ASER 2014 report for explanation of current sample design (ASER Centre, 2015).

Given the architects' goal that ASER would not just capture what children could and could not do, but rather it would provide information that could "feed into strategies of how to improve" schooling, they decided that ASER had to be an annual rapid assessment. ASER would produce test results on children's learning every year, allowing districts and states to track progress and thereby adjust their broader strategies. Dr. Wadhwa's opinion was that one-off data was not very valuable and that it would severely limit the findings' explanatory powers, which cemented Pratham's initial commitment to produce ASER data for five consecutive years. After ASER's frequency was established, Dr. Wadhwa (personal interview, May 15, 2015) articulates that there was a related issue of timing:

> Here, "what time?" was *two* decisions…when would we actually do the survey, and when would the results be available? …. So, we decided that India has similar [school] vacations across the country, children come back to school in July/August, and we *know* that there is always some learning loss: they have been home; they forget. So, that's why every year, ASER is done end of September, October, [and] November. They are back in the new grade; they have had time to settle down; there is some value added in the new class.

With regard to the first timing decision, the architects recognized that in order to measure children's learning tied to their current grade, they would need to test children well after the beginning of the school year to allow children to recover from their summer learning loss and to ensure there was some added value from the children's new classes and teachers. These needs stipulated the decision that testing children in their households would occur September through November. The second decision, pertaining to the timing of ASER results, made it clear to the design team how rapid ASER's assessment and data collection process would need to be, as Dr. Wadhwa (personal interview, May 15, 2015) states:

> And *then*, the other decision…was, when should the results be available? There we decided that we were not going to do it like NSS [the Government of India's National Sample Survey] [that] does a survey and the data is not available or the findings are not available for two years. It takes them two years to crunch the data. We were like no, no, no—two years?…. It is too late! It has to be available [sooner]! We chose and you know

> that we have stuck to that second week of January. The idea was [that]…the government, end of February, is [releasing] the budget every year, so [in] February…[in] the national government, there is nobody available because they are *all* getting ready for the budget. 26th January is a national holiday—it is our Republic Day, and we said, "Okay, we want to bring it out before 26th January, and so if the government wants to use these numbers and feed them into their budgetary exercise, these numbers should be available." So… that second week of January was decided.

Dictating a tight three-to-four-month timeline from testing to publication, Pratham decided ASER findings should be published in time to strengthen the GOI's budget. Highlighting Pratham leadership's close connections with the Planning Commission and general disposition to support the data use of India's economists and finance ministers, Drs. Wadhwa, Banerji and Chavan concluded that the ASER data should guide national budgetary decisions for the education sector.

Innovating a Citizen-Led Approach

With most design decisions made, Dr. Wadhwa clarifies there was still "the final decision" looming: "Who is going to do the survey?" How would Pratham manage to test all these children and collect data in so many households for rural districts throughout India? She describes the challenges confronting them:

> Even though it is a pan-India organization, Pratham does not have that kind of people who can do it…simultaneously across the country. Also, given that we were testing reading and language, we realized that it *had* to be done in regional languages. You can't take people from Delhi, who will go and test in Telegulu in Andhra Pradesh,[9] so it had to be some local [people]. (W. Wadhwa, personal interview, May 15, 2015)

The first challenge Dr. Wadhwa describes is that Pratham did not have the manpower (in sheer number of employees) to carry out the testing concurrently in every state across India. Moreover, Pratham did not have *local staff* in every state, which highlighted the second manpower problem: because of India's linguistic and cultural diversity, ASER surveyors needed to be locals who

[9] Telegulu is a language typically spoken in Andhra Pradesh and Telangana.

could easily interact with villagers, speaking their regional language and being literate enough in that language for testing children in reading. Dr. Wadhwa (personal interview, May 15, 2015) highlights a related dilemma: whether it was ideal for surveyors to come from within or outside Pratham:

> Internally or externally—that was the question. Internally, we couldn't really do it…. Pratham works in many states, but [it] is not that we are doing it in the areas where we work. We wanted the separation between [Pratham programs and ASER]—you know, when you say, "the provider can't be the *regulator*?" The government is the provider [and] if it is also the regulator, then immediately you have vested interests, right? So given that Pratham was a *provider* of solutions, it could not be also [the regulator], so we said, "it has to be external."

From the perspective that ASER is evaluating the quality of schooling or "solutions" (i.e., learning interventions) that address lacking primary education, Pratham leadership saw that both the Indian government's and Pratham's involvement presented a problem of vested interest in ASER. While Pratham undoubtedly would be the financer, coordinator, and supervisor of the ASER data collection efforts as well as the eventual analyzer and publisher of ASER data, Dr. Wadhwa infers that involving external people as surveyors added a layer of transparency and trustworthiness to ASER that would not be there otherwise. However, the question remained, who could fulfill this need for external data collection?

> Now, external could be hiring an external agency to do it. Well, [with] external agencies, one [barrier] is of course the cost constraint. The other was that at some level if you are going to build *awareness*, if the goal is to shift the focus from inputs to outcomes…you have to build that awareness at *all levels*. It is not just at *central* policymaking. It is *state* policymaking. It is parents. It is the *common* man, right? That is where the *citizen-led* kind of nature came out of it. (W. Wadhwa, personal interview, May 15, 2015)

The "citizen-led" nature of ASER emerged not only from a desire for external involvement in data collection within certain cost parameters, but also—and more importantly—from Pratham's aspiration to engage people at all levels of governance and society in the cause of primary education. From the powerful central policymaker to the "common man," the ASER design team

was concerned with how it could build awareness in all of them about the status of primary education as measured by children's learning outcomes. Pratham leadership believed that ASER data would inform policymakers, but what about people at the grassroots? To solve this issue of the need for localized manpower while raising awareness at the ground level, Dr. Wadhwa (personal interview, May 15, 2015) expresses that the design team recognized not only the logistical benefit, but also the power of involving the ordinary citizen directly in data collection:

> As researchers, [we] see many tables, and many tables depending on what you are working in have bad [news]: 50% children in India are *malnutritioned*; 50% of children in India, like ASER says, cannot read a Standard II-level text. Yes, these are shocking numbers. But, what is *far* more shocking is to go into that village and to sit down with that child; you have recorded that the child goes to school; the child is in grade 5, and you see the child *struggling*. Now, *nothing* brings the problem home until you *actually see* that. It's like, you know, it is one thing to talk about malnutrition, but it is another thing to see that child with that *potbelly*….You see Indian kids with blond hair; Indians don't have blond hair. Blond hair is coming from leeching because they are so malnutritioned. When you actually see that, there is nothing that brings the problem home [as that does].

Through her malnutrition analogy, Dr. Wadhwa conveys that the ASER design team understood a dual reality: one of quantitative data and one of firsthand experience. Because they are academics with quantitative data experience, the architects knew for policymakers and researchers the intellectual benefit of numbers that reveal India's difficult reality (e.g., national figures on literacy, malnutrition, and income). Nevertheless, the ASER design team also recognized these numbers' inevitable abstraction of people's everyday struggles. Pratham—from its grassroots staff to its leadership—was (and is to-date) an organization of education fieldworkers with experience in urban slums and remote villages. As fieldworkers, the Pratham leadership comprehended that statistics were intangible, but firsthand experience humanized these issues, this reality that numbers attempted to unveil. In short, they believed if they could involve ordinary citizens in fieldwork as "volunteers," Pratham could build an unparalleled grassroots awareness across India about the status of its primary education system.

Dr. Wadhwa (personal interview, May 15, 2015) reveals that the architects hoped their citizen-led approach would fundamentally change people's level of personal investment in education:

> So, I am giving my *time* to do this, right? I mean we don't *pay* our volunteers. Sure, we pay [i.e., reimburse] their [travel] expenses, but they give their time…. These [choices] are *built* into the architecture of ASER; I mean that citizen-led nature is *built into* the architecture; it is *intrinsic* to ASER…. And, it all came from, what is it that we wanted this thing to *do*?

Through firsthand engagement in testing children and talking to families in villages about schooling, the ASER architects wanted ASER volunteers and villagers to evolve in their sense of investment in children's education and ownership of the Indian public school system. By shifting people's "focus from inputs to outcomes" (W. Wadhwa, personal interview, May 15, 2015), ASER was designed to motivate action at the policy level through its data and at the ground level through its citizen-led design: it asks citizens to donate their time toward traveling to villages in their district and investigating what children could do, believing that such an experience would alter essentially citizens' commitment to increasing what children could do.

Conclusion

"When you began ASER, did you have a sense of what success would look like?" Dr. Chavan (personal interview, May 20, 2015) responds, "I think *doing* the survey was a success…reaching all [of] India…. Pratham was not even *in* all those states, so getting the survey done and getting it done in *time*." Accounts from Pratham staff who participated in that first ASER reveal it was a hectic, exciting, and unsolidified process. During the course of ASER 2005, Pratham workers fanned out across states in search of help to execute it. Drawing on connections Pratham had to civil society organizations, community groups, universities, the private sector, and influential individuals, its workers formed partnerships and recruited a variety

of volunteers to carry out ASER testing. Pratham received an inordinate amount of financial and technical help. For instance, in less than 100 days, 530 partnerships were forged with approximately 20,000 volunteers executing the surveying and testing in 15 different languages across 509 rural districts of India (with 485 districts' data ultimately reported in ASER 2005) (see Pratham Education Foundation, 2006, pp. iv-v, 1-5).

The narratives of ASER's three architects reveal how the two origins of ASER in Pratham programming and Government of India policy conversations substantially shaped its methodological design and goals. Concerns about ASER's usefulness to GOI policymakers governed choices regarding ASER's sampling, frequency, and timing of its results. Knowledge gleaned from Pratham programs guided the architects' choices regarding how to define "learning" in ASER, where to conduct testing (in households or schools), and why simple assessments to gauge children's basic reading and math skills were important to community interest in the ASER process.

Made possible by its simple testing tools, ASER's citizen-led approach is the defining characteristic of its design and has cemented ASER's designation by transnational experts as pioneering in the investigation of children's basic learning (see Pritchett, 2013). Recruiting ordinary citizens to voluntarily conduct ASER ensures its surveyors have the local language and knowledge necessary for effective testing while being able to execute ASER inexpensively and at India's national scale. Also, Pratham sees citizen surveyors as contributing to ASER's transparency and trustworthiness while facilitating widespread engagement in education, thereby increasing everyday awareness of children's learning outcomes. In a country where adults in villages frequently have limited experiences of schooling themselves, ASER's citizen-led design

serves as a platform for altering people's and communities' sense of ownership over the education system and sense of responsibility regarding the trajectory of children's learning.

ASER has both a democratizing and development-focused influence on educational research within India: democratizing in that ASER asserts meaningful information about children's learning can be collected (without educational experts) using ordinary citizens and very simple tests; and, development-focused in the sense that ASER is especially attuned to the needs of planners and economists. ASER's design was not principally aimed at facilitating pedagogical changes at the classroom level, rather its representativeness and timing were considered in relation to planning and the national budget. Nonetheless, ASER data is used to make arguments that have implications for education practice: In public discussions regarding primary education reform, ASER data is used to advocate for adopting Pratham's instructional model of "teaching at the right level," which focuses on children's competency-level versus grade-level in reading and arithmetic instruction (Banerji & Duflo, 2015).

ASER's long-term goals are straightforward. At the policy level, the ASER design team wanted to shift how policymakers viewed the Indian education system's success by focusing on learning outcomes versus inputs. This was accompanied by a related goal of incentivizing and facilitating policymakers' use of educational data in their planning. At the ground level, the architects wanted to make parents aware that schooling did not equal learning. They also wanted to increase ordinary people's sense of responsibility over children's learning and ownership of government schools. These goals constitute the architects' shared vision of what ASER could eventually accomplish. ASER's vision also reflects several of Pratham's articulated priorities, like promoting "programme results and accountability" and being "dedicated to large-scale change" (Pratham Education Foundation, 2016).

"We went in favor of rebellion" infers forging a path not hostile to government's efforts, but rather characterized by independence and the assertion of ordinary citizens' rights in a democracy to evaluate the outcomes of their public systems. The architects envisioned a survey and assessment process that would eventually spur popular action to improve primary schools and children's learning and empower policymakers to use outcome-oriented data in their decision-making. Because its creators viewed these two changes as fundamentally challenging the status quo of how people viewed primary education and how policymakers' made decisions, ASER felt rebellious in its quest to help the "People of India, From different states and regions, Speaking different languages…. To build a better India" (Pratham Education Foundation, 2006).

Postscript

This paper recounts pieces of ASER's creation story from the perspective of its architects. In the M&E sphere of large-scale data where ASER resides, the prevailing view may be that narratives are not valuable because unlike numbers, stories are subjective. Stories have a point of view: they are told from (and therefore limited by) a particular social standpoint. However, narratives are not untrue—or unworthy sources of data—just because they cannot easily be disentangled from the perspectives of their storytellers. All forms of data have a point of view, which is an important truth when investigating the design of monitoring and evaluation efforts: quantitative data are produced through judgments of what to measure or study and how, of what to report to an audience and why (Clarke & Dawson, 1999).[10] Such data, once collected, once shared, are in turn interpreted by others as meaningful, valuable, credible—or not. Through this paper, I have endeavored to lay bare the story of ASER's design, which includes its data's

[10] Clarke and Dawson (1999) highlight that "quantitative indicators" used in monitoring and evaluation efforts are not "entirely objective and value free"; such data also require a "critical evaluator" to ask questions about the data's point of view, such as who defined the "indicators in the first place and for what purpose" (p. 6)?

point of view, but this reality of ASER's subjectivity does not make its data inherently less valuable. It is helpful to consider how rarely the social and political origins of large-scale data are studied. And, to further ask: How do these origins, in turn, shape our judgment of large-scale data's social meaning and political consequences?

PAPER 2. Theorizing from the Ground: Ethnography of a Citizen-Engaged Evaluation Process in India

Abstract: The *Annual Status of Education Report* (ASER) evaluates the condition of primary schools and children's basic learning in rural India. ASER is not only a large-scale technical endeavor but also a far-reaching social process. One can analyze ASER as a unique form of evaluation—a model for *citizen-engaged evaluation*—that investigates India's progress in universalizing primary education by innovatively engaging parents, teachers, volunteers, and members of partner organizations in its data collection process. This paper draws upon the academic literature on evaluation in order to examine aspects of ASER's model and process as well as explore its potential as "a kind of social conscience" (Schwandt & Gates, 2016, p. 67). Furthermore, because ASER is an evaluation developed outside of the Global North, this study posits how theorizing about ASER from the ground may enrich the field's literature and expand the appropriateness of evaluation concepts for non-Western contexts.

> *"When ordinary people learn to measure what affects their lives, they can communicate with each other across villages, states, nations, and continents, to identify and understand their problems, take steps to resolve them, and change the world for the better"* –ASER Centre, n.d.b, vision statement

The *Annual Status of Education Report* (ASER) investigates the conditions of primary education in all rural districts of India. Every year, ASER Centre—an Indian nongovernmental organization headquartered in New Delhi with satellite offices in states across the country—conducts ASER, independent of the Government of India (GOI).[11] The report is produced from data collected in villages by volunteers with the help of ASER Centre's partner organizations. ASER volunteers perform village asset mapping, government school surveys, and household-based surveys and learning tests to generate the report's data. The report comprehensively maps, by state and district, Indian children's ability to read simple texts and do basic math. These

[11] ASER has been conducted every year since 2005, with the exception of 2015 when it was suspended to allow ASER Centre leadership and staff to pursue pilot research related to expanding what ASER currently measures through its assessments and surveys.

reading and math tests in ASER provide the only annual and publicly available figures on children's learning levels in India.[12] Moreover, because learning tests and surveys are administered in children's homes, ASER learning and education-related data (e.g., enrollment, parental education levels, and household affluence) includes children who may typically be missing from other large-scale educational data sets in India; for instance, ASER data includes children who attend madrasas (i.e., Islamic religious schools) and unrecognized private schools, who are frequently absent from school, and who never attended or dropped out of school (Goodnight & Bobde, unpublished). The scope of ASER far exceeds even the Government of India's National Sample Survey (NSS), which is its "main source of data for estimating poverty, employment and for other socioeconomic indicators" (ASER Center, 2015, p. 69). For instance, NSS (2011-2012) sampled a total of 59,129 rural households whereas ASER 2014 surveyed 341,070 rural households in 16,497 villages across 577 districts of India. In addition, ASER administers learning tests to approximately 650,000 children ages 5-16 years old annually. Remarkably, despite ASER's enormous scope and growing reputation within India and abroad (e.g., Barrett, 2011; Jha & Parvati, 2014; Learning Metrics Task Force, 2013; Nawani, 2013; Pritchett, 2013; Results for Development, 2015), the process underlying ASER has been under studied.

Unlike the testing and surveying efforts coordinated by the governments of India and other countries, ASER presents a unique model for the evaluation of social systems due to its independence, household-based approach, capacity-building components, and inclusion of

[12] ASER is designed to produce representative estimates at the district-level. In 2009, Ramaswami and Wadhwa conducted a study that looked at the precision of ASER state and district level estimates, which found that whereas state level averages were within "a margin of error of 5% or less," district level estimates were "less precisely estimated" because of ASER's second stage of clustering (i.e., household selection) "increases the variability of estimates" (ASER Centre, 2015, p. 275). Thus, ASER 2014 reports divisional estimates for districts of most states, but not for some smaller states with fewer districts (and smaller populations of children within districts) like Manipur.

volunteers and partners.[13] I refer to this pioneering form of evaluation as *citizen-engaged evaluation* because it relies upon the participation of hundreds of partner organizations and thousands of volunteers. Furthermore, ASER is an evaluation model that champions ideas about civic participation: its partners and volunteers are organized under the principle that ordinary citizens have rights within a democratic society to independently evaluate their public systems. Thus, the "citizen-engaged" aspect of ASER is born not just out of a logistical need to overcome the data collection issues of scale, cultural diversity, security, and cost, but also out of a philosophical stance on popular engagement in education (Goodnight, Paper 1). With overtones of government accountability and civic responsibility, ASER Centre's vision statement further illuminates the philosophy animating ASER: "When ordinary people learn to measure what affects their lives, they can communicate with each other across villages, states, nations, and continents, to identify and understand their problems, take steps to resolve them, and change the world for the better" (ASER Centre, n.d.b). This vision is one of self-determination and public ownership of the systems and institutions that affect people's lives. Linking community-level agency to national (or global) action, the vision also implicitly speaks to evaluation's potential to encourage cooperation, knowledge sharing, and collective problem-solving for increasing the public good.

According to the leadership of ASER Centre (and the academic literature produced on ASER to date), ASER is a "citizen-led assessment" because the people conducting the testing of children's learning are volunteers supported by local partner organizations, but I argue this language obscures more nuanced interpretations of ASER's structure and process. In reality, the

[13] When I refer to ASER's evaluation "model," it comprises ASER's methodological design and entire process for training, data collection, monitoring as well as the data subsequently produced from its surveys and tests and the report itself. Also, included in this model is the Centre's routinized process for the dissemination of ASER findings, i.e., its promotion of every annual report with the media and its outreach efforts to government, academia and civil society organizations.

Centre directs the full ASER process that includes designing surveys and tests, training, sampling, monitoring, reporting and disseminating findings—this suggests that ASER, instead of being "citizen-led" is actually citizen-engaged. ASER is facilitated (not led) by its partners and volunteers who are, nonetheless, undoubtedly crucial to its execution and ultimate success. As the paper progresses, I will argue that it is appropriate to conceptualize ASER beyond its main research method for generating data: assessment. Instead, ASER should be conceived more fully as an evaluation because 1) the information generated by its other methods (i.e., field visits, surveying, and mapping) is equally meaningful and 2) the goals of ASER are much broader than simply reporting the learning levels of children gleaned from tests. Its report is making an evaluation of India's primary education system in order to prompt public dialogue about the status of the nation's universalization efforts and issues with the education system's governance (Goodnight, Paper 1).[14] Through collecting and disseminating evidence, the Centre seeks to generate debate on the effectiveness of schools (as judged in part by children's basic learning levels) under India's current policy, the Right to Education Act (RTE) of 2009 (Government of India, 2009; see Goodnight, Paper 1). The architects of ASER believe such public dialogue based on quantitative evidence can create the political pressure necessary to motivate key policy changes that improve the quality of primary education.

This paper explores how ASER Centre's vision of "ordinary people learn[ing] to measure what affects their lives" shapes its evaluation process (ASER Centre, n.d.b). The paper includes seven sections: following the introduction, the next section reviews evaluation literature, highlighting definitions and qualities of evaluation while exploring how evaluation relates to the concept of monitoring in non-Western countries. Afterward, I outline the study's questions and

[14] Some evaluation experts may argue that ASER's model is more accurately described as monitoring versus evaluation. I will address this further in the next section.

its methods and data. The fourth and fifth sections offer a brief history of ASER's development and then, an extended discussion of ASER's evaluation process, particularly its recruitment, training, and data collection processes. In the discussion section that follows, I examine ASER according to the definitions and qualities of evaluation that were previously identified (in the review of evaluation literature). The paper concludes with revisiting notions of ASER as a citizen-engaged evaluation and exploring how the example of ASER may expand evaluation theory and practices in the Global South.

Defining Evaluation and Identifying Key Concepts Across Contexts

> *"We seek an evaluation practice [that serves] as a critical voice in social, political, and cultural environments indelibly marked by significant inequalities, power differentials, uncertainty, ambiguity, and interpretability. This is an environment in which we ask difficult questions about what kind of society we should have, what directions we should take and in which we are, at best, able to give only partial and temporary answers. Evaluation in such an environment is a kind of social conscience. It involves serious questions of public direction. It is a risky undertaking in which we endeavor to find out not only whether what we are doing is a good thing to do but also what we do not know about what we are doing." –Schwandt and Gates (2016, p. 67)*

There is genuine debate over what evaluation is, and is not. Schwandt and Gates's (2016) definition offers a strong moral vision for what evaluations can aspire to be amidst circumstances that are "marked by significant inequalities, power differentials, uncertainty, ambiguity, and interpretability" (p. 67). Their vision suggests that evaluations can offer hopefulness and a direction for collective action in situations where the barriers to social welfare are many and the challenges to social policy and programs seem overwhelming. In this definition, the cold technical practice of evaluation meets the warmth of humanitarianism and the messiness of unequal relations, discordant lived experiences, and uncertain futures.

Depending upon how evaluation is actually practiced and for what purposes, it can be any number of things, and it can fail to be them. I find that short and simple articulations of what

evaluation *is* can characterize its basic essence in ways that are intellectually potent. For instance, evaluation is "a kind of social conscience" and a "critical voice" (Schwandt & Gates, 2016, p. 67). Evaluation is about "cultivating the critical voice of others" (Schwandt & Gates, 2016, p. 68). Evaluation is "determining" the "worth" of something (Scriven, 2013, p. 170). Evaluation is "the systematic collection of information" (Patton, 1982, p. 35). Evaluations are "arguments in which evaluators [present] evidence for and against" (House, 2013, p. 200). In sum, "'evaluation' can be described as 'an elastic word that stretches to cover judgments of many kinds'" (Weiss, 1972, p. 1, as cited in Clarke & Dawson, 1999, p. 1). In later sections, I will revisit Schwandt and Gates's (2016) long definition as well as these shorter articulations of evaluation to discuss features of ASER.

As the academic literature on evaluation grows so does the difficulty with defining the parameters of *evaluation* that make it distinguishable from other forms of research (Alkin, 2011; Schwandt, 2009). As mentioned, Michael Scriven (2013) equates evaluation with the act of valuing itself—evaluation is "the process of determining…something about the merit, worth or significance of any entity" (p. 170). However, "valuing" alone is an unsatisfying criterion in distinguishing evaluation from other social research (Mathison, 2014): valuing can be a component of many, quite dissimilar research endeavors.[15] Meanwhile, Marvin Alkin (2013) writes about the intense context-sensitivity of evaluation as a "major distinction between 'research' and 'evaluation'"—a "distinction [that] rests heavily on the difference between producing generalizable knowledge and providing information that will be helpful in improving specific programs" (p. 283).[16] Nevertheless, this binary between (evaluation's) helpful

[15] It would seem that a substantial amount of research on education—whether it is on curriculum, teaching strategies, or school culture—has a valuing component within it.

[16] The evaluation-research binary suggested by Alkin (2013) reflects a dualistic way of thinking that assumes other dichotomies exist as well. For example, his description of "evaluation" versus "research" is

information for improving programs versus (research's) generalizable knowledge for the greater world is subverted within a multi-disciplinary, practice-oriented field such as comparative education.[17] Elsewhere, Alkin (2011) acknowledges that "sometimes the two are virtually indistinguishable" as research and evaluation "are [both] forms of disciplined inquiry," but he posits, "the former *seeks conclusions* and the latter *leads to decisions*" (p. 8). While affirming their respective "uniqueness," Mertens (2009) reveals that evaluation and research have noteworthy intersections or "shared territory" especially with regard to advocacy and social justice (p. 2). In total, the features of evaluation can overlap greatly with those of social research—the parameters drawn around evaluation to separate it from other forms of research are somewhat faint; nevertheless, attempts to define and highlight evaluation's uniqueness from research have been important to constituting 1) the discipline of evaluation, 2) a professional evaluator identity, and 3) an academic literature that informs the design, practice, and impact of evaluations (Christie & Alkin, 2013; Mathison, 2014).

Evaluation's definitional specificity and concrete parameters are sabotaged by its range of possible applications and purposes. Evaluation theorizing and guidelines for practice are further complicated by the field's foundation in multiple disciplines and epistemologies (Mathison, 2014). However, the evaluation field consequently enjoys considerable methodological flexibility and theoretical breadth (Christie & Alkin, 2013), making evaluation

---

predicated on a local-global dichotomy (that indicates a small-scale versus large-scale inquiry project with the intended outcome of context-sensitive findings versus generalizable claims. Also, Alkin asserts there is an information-knowledge binary wherein evaluation and research have divergent goals of either prescribing practice or generating theory. In other words, there exists a dichotomy of identifying facts or explaining truths within a specific environment (i.e., information) versus discovering tenets that apply across space and time (i.e., knowledge).

[17] Some comparative education studies are ethnographic, context-driven, concerned principally with meaning-making amongst a small group of people, and focused on the shared values, norms, or culture within a specific place (e.g., community, school, or organization). Comparative research can be done in the spirit of trying to produce findings that will help schools, communities or organizations to function better or improve their practices—such research can also be concerned with the influence of its findings on the people with whom the research is done. Also, qualitative research, and particularly ethnography, often avoids claims that the knowledge it produces is "generalizable" (see Maxwell, 2013, pp. 136-137).

an adaptable and potentially useful form of inquiry in any number of areas and contexts. It is this plurality—in discipline, epistemology, methodology, and application—that generates spirited debate about evaluation today.

Though evaluation theorists have outlined several types of evaluation and purposes for it (Greene, 2016), the following are significant questions for analyzing a broad range of evaluations:[18]

- Whose interests does the evaluation serve? What purpose for the evaluation is stated? What goals are promoted?
- What (and whose) questions are pursued and not pursued through the evaluation?
- What role(s) is/are prescribed for and fulfilled by the evaluator(s)?
- What roles are prescribed for and fulfilled by stakeholders?
- How inclusive is the evaluation process of the various stakeholders who are invested in the entity being evaluated?
- What are credible data for judging the entity's value in the eyes of various stakeholders?
- Who is/are the intended audience(s) for evaluation findings and recommendations?

In evaluations that are intended to be "transformative" (Mertens, 2009), "democratic" (House & Howe, 2000), and "equity-focused" (Greene, 2016), a main concern is the inclusiveness of typically underserved stakeholders and the quality of these stakeholders' participation during the course of the evaluation as well as their degree of influence over the questions the evaluation pursues. Relatedly, there is a need for evaluators to be reflective about how accessible and useful

[18] Alkin (2011) identifies four kinds of evaluation defined by the thing that is being evaluated: "*products*, *personnel* (or individuals), *policy*, and *programs*" (p. 7). In his book on the essentials of evaluation and in his writing more broadly, Alkin focuses—as the majority of evaluation theorists do—on the evaluation of "programs," which have no singular definition across the field. Large-scale evaluations of social systems—like education systems—seemingly fall conceptually somewhere between policy and program evaluation.

the evaluation's findings are to all stakeholders and whether the data collected on the *evaluand*, the evaluated entity (i.e. program, policy, or system), is credible.

Monitoring and Evaluation

Monitoring can be defined as a form of inquiry that is distinct from evaluation, or it can be used as a synonym for or in unison with evaluation (i.e., "monitoring and evaluation"). Some evaluation experts argue that when evaluation is principally for accountability purposes, it is monitoring—also known as "performance monitoring" or "performance measurement" (Ryan & Cousins, 2009, p. xii; Ryan & Feller, 2009, pp. 173,176). Thus, one difference between evaluation generally and monitoring specifically appears to be a focus on accountability. "Divergent evaluation purposes" (Ryan & Cousins, 2009, p. xii) and distinctions in scale are reasons why monitoring and evaluation have been commonly conceived of as separate things.[19] For instance, Gildemyn (2004) perceives monitoring and evaluation as "two distinct but complementary processes" with "twin goals" of "accountability and feedback/learning" (p. 522). Amongst evaluation theorists, monitoring has been viewed chiefly as large-scale, concerned with quantitative outcomes, and unresponsive to contextual differences whereas evaluation (i.e., program evaluation) is locally specific, concerned with qualitative processes, and aimed at learning and improvement.[20] There is growing concern that evaluations, which are context-sensitive or "responsive" to the needs of stakeholders, are being "supplemented or supplanted by

[19] Articulating commonplace ideas regarding the differences between monitoring and evaluation, Ryan and Cousins (2009) state: "Single, stand-alone educational evaluations… aimed at learning and discovery at the local program level are under some duress—either being supplemented or supplanted by outcome indicators (Dahler-Larsen, 2006; Mayne & Rist, 2006). Meanwhile, despite known conceptual complexities (e.g., divergent evaluation purposes), performance monitoring is identified as one evaluation approach aimed at practice improvement and organization learning (Rogers & Williams, 2006). Although performance monitoring's historical role in holding individuals and organizations accountable is recognized, performance monitoring is becoming known as a means to improve programs, services, and practices, effectively blurring foundational evaluation purposes" (p. xii).

[20] Accordingly, recent models for teacher evaluation that are system-wide and use indicators like students' test scores would actually not be considered evaluation in the traditional sense but performance monitoring or "appraisal" according to Schwandt (2009) and Alkin (2011).

outcome indicators" (Ryan & Cousins, 2009, p. xii), inferring that indicators distort evaluation's focus on improvement and mark an ominous shift toward an expanded use of indiscriminate, overly "technocratic" monitoring (Chouinard, 2013). With the steep growth in monitoring and evaluation (M&E) activities worldwide, it appears some evaluation theorists have anxiety around dissolving distinctions between the two, particularly given the myriad purposes for which monitoring is being used.[21]

A field's shared terminology and language facilitate collective understanding of its concepts amongst its practitioners and theorists. While recognizing its essential role in academic fields and disciplines, terminology is also necessarily debated, subverted, and rejected. At times, policing terminology (especially, that which leads to binary thinking) is misplaced vigilance that restrains new understanding instead of facilitating it. The next section highlights how terminology across contexts can be one barrier in identifying and further strengthening evaluation. As evaluators, researchers, policymakers, and practitioners confront the realities of why evaluations are mushrooming across the world, they face the challenges and opportunities greater global interest in evaluation presents. A fluid perspective of monitoring and evaluation is helpful for ascertaining how M&E projects are actually being implemented, understanding what influence they are having, and learning from which constraints they are surmounting through innovation. As such, with respect to this study, monitoring and evaluation is treated as conceptually adjoined and M&E is discussed more generally under the term *evaluation*.[22]

[21] See footnote 9.

[22] In judging the merit of an evaluation, this study posits that of greatest importance is judging the evaluation's purpose and the appropriateness of its methods and data for achieving that purpose. Closely analyzing an evaluation's components—its purpose, goals, questions, scale, methods and data—appears more immediately fruitful than guarding semantic distinctions.

Evaluation and Terminology across Contexts

Because issues of context matter greatly in evaluation (Alkin, 2013), the field has been grappling more recently with the Western bias in its theorizing—that is, in its conceptualization of evaluation, formulation of evaluation ethics, and establishment of guidelines for evaluation practice. The normalization of North American contexts within the field's literature has restricted its appropriateness and usefulness for nations outside the United States, Canada, and Western Europe.[23] In countries from the Global South—often referred to as "developing countries,"[24] what evaluations ought to address, at times, is different from the challenges and goals typical of "conventional evaluations," which are designed for high-income Global North nations, as Ofir and Kumar (2013) explain:

> There is a growing feeling, especially in developing countries, that most evaluations are neither useful nor responsive to the needs of policy makers and other stakeholders. For start, many conventional evaluations, by focusing too narrowly on a particular project, end up accounting for inputs and a few scattered outputs, outcomes, or impacts.... Knowledge and lessons learned are seldom synthesized and presented in a manner that can be used beyond a particular project or program.... Part of the explanation can be found in the failure of conventional evaluation methods to take note of both the needs and circumstances of developing countries. Developing countries have important characteristics that should shape development evaluation practice, and yet they are often

[23] For example, Cousins's (2015) glowing *American Journal of Evaluation* book review of Marv Alkin's (2011) much celebrated text, *Evaluation Essentials: From A to Z,* had only one critique: its "Western, if not North American-centric" orientation. Cousins questioned the book's "applicability in cross-cultural contexts" but admitted he does not "have a good answer to that one." Cousins's observation came, in part, from using the text as a handbook with new evaluators in India where he was involved in evaluation capacity-building efforts and "a nationwide evaluation of secondary teacher in-service training" (p. 418).

[24] Though "developing" is a label still in common use, I feel it reproduces a colonialist, white supremacist conceptualization of countries in the non-white world. Unfortunately, the use of this word can reinforce a misguided view, which perceives sameness between diverse contexts because of beliefs about their common backwardness, deficiency, poverty, or dysfunction along a predetermined civilization or modernization trajectory. Accordingly, nations (or their citizens) are seen as developmental levels behind the science, rationality, and advances of Europe and North America. In short, adjectives like "developing" have unwittingly racist and social Darwinist underpinnings. With that said, countries can have similar experiences of first-world oppression and their respective conditions of postcoloniality may allow for studying fruitful connections between them. I use terms like the "Global South" or "postcolonial nations" in an attempt to avoid the white supremacist tinge of this older term while noting that experiences of colonization, imperialism, neoliberalism, and general economic, cultural, and military domination via a legacy of development regimes (like structural adjustment policies) have forged similarities between nations typically labeled "developing" where commonalities between them might otherwise not exist.

> ignored or underplayed in methodology guidelines, to the detriment of credible and useful evaluation. (pp. 11-12)

Ofir and Kumar (2013) highlight many tensions between, on the one hand, how evaluation is normally construed and practiced and, on the other hand, the needs within non-Western nations. The following are four substantive tensions that they indicate in their text:

1. **Useful information.** One tension is whether evaluations produce information that is useful to policymakers and other stakeholders in the Global South. This requires the evaluation's purpose to address stakeholders' needs as *they* identify them—i.e., being "responsive."
2. **Credible data.** Related to this issue of purpose, a second tension is whether the indicators or types of data being collected are viewed as credible measures of the entity under evaluation. Alternatively, the indicators can be irrelevant "inputs and a few scattered outputs, outcomes, or impacts" (Ofir & Kumar, 2013, pp. 11), which do not credibly represent the program or system according to perspectives within that country.
3. **Relevant scale.** A third tension is whether evaluations, instead of "focusing too narrowly on a particular project" (Ofir & Kumar, 2013, pp. 11), are able to tackle the scale at which decisions often need to be made (and for which policy is often created): state and national systems.[25] Evaluation theorists commonly discuss ethics, principles, and practices of evaluation in relation to the narrower circumstances of evaluating "programs," but in the above quotation, Ofir and Kumar (2013) observe policymakers regularly want information "beyond a particular project or program" (pp. 11). As such,

[25] Many countries in the Global South have large social schemes and systems under increased scrutiny as a result of the Millennium Development Goals (MDGs) and now, the Sustainable Development Goals (SDGs). The MDGs sparked an era in which comparative, worldwide data on social indicators was seen as important for coordinated target setting and monitoring the progress toward national and international goals. The SDGs have continued to heighten the pressure on nations to use data effectively to understand their progress in areas such as basic education.

evaluation literature may have limited value in designing large-scale evaluations, which are in high demand. Scale greatly affects the overall evaluation approach and more specifically, the choice of methods and the level and nature of stakeholder participation.[26]

4. **Meaningful presentation.** A fourth tension noted by Ofir and Kumar (2013) is the manner in which evaluation findings or "knowledge" is presented—information is of minimal consequence if not "synthesized" or articulated in a way that is accessible and can be applied toward decision-making or enabling stakeholders to take action (pp. 11).

In India and in many postcolonial countries, the Western-centric perspectives and values embedded within mainstream evaluation theories necessitate substantial *theory translation* (Goodnight, 2017), which goes beyond logistical or practical modifications to ways of seeing and valuing.[27] Ofir and Kumar (2013) observe: "Attention is seldom given to the broader frameworks or sets of values that determine the focus" of evaluations in the Global South (p. 12). Moving beyond the unreflective application of Western values in M&E is a persisting issue of global justice. As Ofir & Kumar (2013) remind us, postcolonial countries cope with a "world dominated by Euro-American perspectives, models, and power…. [that] tend to induce or exacerbate feelings of inadequacy, disempowerment, and marginalization," and the consequential "effects on the psyche of an individual, institution, or nation are often underestimated" (p. 15). The logical follow-on from this observation is, how does the field move beyond the default Western orientation of its theorizing?

[26] Resources like money and institutional capacity, which vary widely across nations and sub-national contexts in countries like India, also influence decisions regarding scale, methods, and stakeholder participation for evaluations.

[27] Elsewhere, I have defined *theory translation* as "the act of translating a theoretical framework from its original context to another place for the purpose of conducting and interpreting social research" (Goodnight, 2017, p. 1), and I have noted the strong ethical and philosophical considerations in this process, especially when the resulting theory will be used to determine the value (e.g., equity or quality) of a social system, as in India's education system.

I identify two approaches for coping with the current theoretical limitations of the field, which are theory translation (as mentioned above) versus theorizing from the ground. In 2002, Nick Smith, recognizing the "cultural embeddedness" of U.S. ideals in evaluation approaches, solicited his international students' "reflections" on evaluation theory because he thought they eventually "must…attempt to translate the theory to their own contexts, adjusting for differences in culture and socio-political realities, and assess its possible utility in their own evaluation work back home" (p. 481). Fifteen years later, Smith's (2002) claim is still true: "Locally adapted and culturally sensitive evaluation approaches are needed throughout the world" (p. 481). One solution illuminated through Smith's article is training future evaluators from all over the world in (Euro-American) university programs for evaluation. Another solution is to conduct research on evaluation in non-Western contexts to help generate evaluation theories and guidelines for practice from these evaluation activities in order to produce an expanded field literature that is reflective of the values, norms, cultures, and needs of diverse stakeholder communities. Today, there are evaluations taking place in numerous countries, but they are not necessarily being documented or researched in ways that inform the field's theories.

A practical challenge in studying evaluation across contexts is terminology—*assessment*, *appraisal,* and *evaluation* are terms that evaluators have been unsuccessful in policing (Schwandt, 2009, p. 19). At times, researchers and practitioners, operating between the fields of evaluation and education, have used words like *assessment* and *evaluation* interchangeably (Alkin, 2011, p. 10).[28] Various terms for evaluative activities proliferate in contexts like India that generally lack a nation-specific literature on evaluation. A. K. Shiva Kumar (2010) observes:

[28] Alkin (2011) defines evaluation, assessment, and appraisal differently—"*Evaluation* is the favored term when we talk of judging a program…. We *evaluate* programs; we *assess* client knowledge [or students]; and we *appraise* staff" (p. 10).

> Many [professionals carrying out evaluations in South Asia] are unable to distinguish between a study, a review, and an evaluation…. Evaluations typically get judged as "good" or "bad" on the basis of statistical rigor—not recognizing that a good evaluation is not the same thing as a well-designed survey. (p. 283)

As an example, India's one-off, large-scale research projects in education typically end in a report for funders, government officials, international policymakers, civil society organizations, and/or academics. The inquiry work underlying these reports is called a "survey" or "assessment" even though the project itself requires a range of activities (e.g., instrument development and piloting; field visits or observations; administration of tests, questionnaires, or surveys; data analysis; and report writing and dissemination). The amount and complexity of activities required by such reports can be obscured and minimized by analogizing them to simple methods (e.g., assessment). Furthermore, surveys and tests for these reports are sometimes administered in-person and orally, necessitating travel to educational sites and interaction between researchers and a range of stakeholders (e.g., students, teachers, headmasters, and community members). Through quantitative indicators and sometimes qualitative observations, these reports include findings that are used to make evaluative judgments—i.e., raise a "critical voice"—about the quality or equity of entities like government anti-poverty programs or the public primary education system (see De, Khera, Samson, & Kumar, 2011). Thus, in India, evaluations are sometimes synonymized with methods of research they employ, which I argue is the case with ASER. This habit potentially limits and muddies conceptualizations of what evaluation is or can be. It thereby restricts understanding of how this evaluation within India relates to (and expands) the Western field of evaluation.

Kumar (2010) asserts that "South Asian nations are beginning to increasingly recognize the need for evaluations; but most lack the necessary capacity and competencies…to conduct evaluations" (p. 238). In parts of rural India, there are currently few institutions that may serve

M&E purposes, at least, as typically conceived by "the discipline of evaluation" (Kumar, 2010, p. 238). Yet, the potential M&E capacities that do exist within rural community-based organizations (CBOs) and NGOs in India might be underestimated and thus not nurtured. ASER prompts those of us in the field of evaluation to take a harder look at how we conceptualize what evaluation is and thus, what are the necessary capacities and approaches for conducting evaluations. Relatedly, ASER may challenge or expand our understanding of who can be an evaluator. Given its geopolitical significance, its population's diversity, and the scale of its civil society and government, India is an important context for research on evaluation—especially with respect to expanding current evaluation theories and practices. There is presently minimal documentation on how evaluations are conducted in India (Hay, 2010). As such, it is important to investigate what influence evaluations in India have on decision-making, and in the interim, which approaches they employ for producing data on the quality and equity of social systems.

Evaluation Definition Revisited

This study operates in a somewhat paradoxical state: on one hand, I have outlined the Western biases and limitations within evaluation theories, concepts, and terms, and on the other hand, I utilize some of those same concepts and terms in this study—albeit, cautiously—because I think they can help me make sense of and discuss ASER and perhaps, also challenge some of the field's certainties. Thus, I move forward tentatively defining evaluation as "a form of applied social research" (Clarke & Dawson, 1999), which seeks to provide a diversity of stakeholders like policymakers, educators, researchers, civil society workers, and the general public with credible and useful information; this information about the "social world" is, ideally, leveraged toward the improvement of programs, policies, or social systems (Christie & Alkin, 2013, p. 16) as well as toward the betterment of society, more generally (Henry & Mark, 2003). When

evaluation is used for investigating the quality or efficacy (i.e., value) of large social systems versus smaller programs, I believe the intended goals of evaluation often tilt toward improved policy and governance. However, it is important that such large-scale evaluations do not overlook issues of inclusion, social justice, and equity, which are consistently highlighted as standards for the ethical evaluation of programs.

Following the sections that explain the design of ASER and describe the ASER process, I will introduce and discuss four evaluation concepts: 1) data credibility and usefulness, 2) social accountability, 3) evaluative thinking, and 4) social conscience. These concepts will structure my analysis of ASER in the discussion section. First, however, I will share the research questions, methods, and data that compose this study.

Research Questions, Methods, & Data

A 10-month ethnographic study of the entire ASER process was conducted with prolonged fieldwork in three states: Rajasthan, Manipur, and Tamil Nadu. Drawing on the larger study, this paper explores the following research questions:

1. Does ASER address some of the weaknesses of conventional evaluations in its design and process? How?
2. In what ways does ASER fulfill conventional notions of evaluation and alternatively, challenge or expand them?
3. How can the phases of ASER be understood through evaluation concepts (e.g., social accountability and evaluative thinking)?

To address the research questions, I analyzed ASER 2014 documents, such as training materials and instruction manuals, survey and testing forms, and the published report and findings. I also analyzed field notes from observations of ASER data collection in villages and

from ASER training workshops at the national, state, and district levels. Transcripts from my semi-structured interviews with the ASER architects and ASER Centre directors were also analyzed. Where needed, I consulted other information from web pages and linked documents on the ASER Centre website (http://www.asercentre.org/). Finally, I analyzed ASER 2014 blog posts written in English by ASER central and state team members and partner organizations/MTs.[29] Analysis of documents, interviews, and field notes were guided by a narrative inquiry approach (see Appendix A for more information).

The blog posts represent a unique form of storytelling and particular type of meaning-making engaged in by different participants in the ASER process. These posts reflect more immediate reactions to experiences related to the ASER process than many of the interviews that I conducted with ASER participants. At times, the posts may romanticize the activities comprising ASER given these posts are for a public audience—that audience is mostly of peers, the ASER Centre directors, and supporters of the organization (e.g., funders). Blog posts are examples of the narratives that ASER workers tell each other about their experiences and how they think about the meaning of the ASER process that they all undergo—posts have both individual and collective elements. Using these blog posts as a source of data means I am accessing a selection of highlighted experiences, which ASER participants seemingly believe are most worthy of retelling. Posts are likely not representative of everyone's experiences or even reflective of that blogger's *general* experience—rather, posts likely recount the experiences or moments that mean the most or have the greatest emotional or intellectual impact. The blog posts are good examples of the kind of enthusiasm, sense of discovery, and shifts in perception that

[29] Of the 108 blog posts for ASER 2014, 81 blogs were written in English (and analyzed) and 27 in Hindi (and not analyzed).

ASER generates. The experiences and inexperience of ASER novices are articulated graphically and excitedly through these blogs.

The Development of ASER

Pratham, India's largest educational non-governmental organization (NGO) created and piloted the first ASER in the winter of 2004-2005. Since 1995, Pratham has provided programming in cooperation with government schools and local officials to strengthen the basic education of children throughout India. During the course of its programming in the late 1990s and early 2000s, Pratham workers discovered that children appeared to not be learning at expected levels despite their increasing levels of enrollment and regular attendance in school. In fact, Pratham staff members and program volunteers noticed children "having difficulty with the basics," such as reading simple texts and completing basic math computations, after attending primary school for multiple years (R. Banerji, personal interview, May 20, 2015). This was disheartening and puzzling for Pratham workers, but it was unclear how widespread the issue was—that is, if it extended beyond the areas in which Pratham worked and constituted a nationwide learning problem.

As the NGO was gaining on-the-ground experience in basic educational programming, Pratham leadership was engaged in national policymaking (see Goodnight, Paper 1). In time, the organization's leaders cultivated a unique perspective on the learning and primary school situation in India with a dual focus on 1) schooling at the grassroots level (e.g., in villages) and 2) policymaking, governance, and data use at the national level. Pratham saw a need for annual, national data that could be disaggregated to the state- and district-levels and that would provide systematic evidence of whether children were going to school and whether they were learning foundational skills in reading and arithmetic. Out of these activities and identified needs, the

three architects of ASER, Drs. Madhav Chavan, Rukmini Banerji, and Wilima Wadhwa crafted ASER's trailblazing methodology (Goodnight, Paper 1). After unsuccessful attempts to garner financial and logistical support from the central government, the leadership of Pratham undertook ASER independently. As Dr. Chavan recounts it, "we went in favor of rebellion" (Goodnight, Paper 1, p. 19). The architects believed ASER could serve as a "proof of concept" to government officials, who would be persuaded eventually that they could reliably and inexpensively collect educational data to inform policymaking, budgetary decisions, and school initiatives (Goodnight, Paper 1, p. 18).

In 2005, the central goal of ASER was to find out if children were learning basic skills in reading and math as well as to evaluate those learning levels in context of information collected about their enrollment and household. ASER also collected data about the basic condition and functioning of government schools.[30] When ASER began, Pratham leadership anticipated conducting it annually for five years, but ASER was extended beyond that period after recognizing the continued need for yearly data on government schools and children's basic learning levels. (Also, the Government of India and state governments failed to take up efforts to collect annual data on basic learning despite ASER's "proof" that the government could manage it.) In 2008, Pratham was awarded funding from Google to establish an independent NGO, called ASER Centre, which would assume responsibilities for conducting ASER.[31]

After the passage of India's Right to Education Act (RTE) in 2009 (Government of India, 2009), the school surveys administered during ASER included specific indicators to evaluate RTE's implementation:

[30] Schools visits and surveys were conducted in all years of ASER except 2006 and 2008 (see "From 2005-2014: Evolution of ASER" in ASER Centre, 2015, pp. 66-67).

[31] See http://www.asercentre.org/ for more information about ASER Centre (e.g., its mission, work, and history).

> This year ASER collected data on those RTE norms for which compliance can be easily observed, during the school visit. The RTE specifies clear norms for enrollment, access, school infrastructure, teacher appointment, TLM [teaching and learning materials] and pupil teacher ratio (PTR). Most of these are easily observable, or data can be collected to check if they are being adhered to. However, where the RTE norms are fuzzy is in the area of children's learning achievement. Phrases like "building up child's knowledge, potentiality and talent" and "development of physical and mental abilities to the fullest extent" are used. In many ways, the RTE continues the tradition of focusing on inputs rather than outcomes. (ASER Centre, 2011, p. 8)

The passage of RTE provided ASER Centre with further indication as to what measures the GOI would see as credible indicators of school quality (see school survey form in ASER Centre, 2015, pp. 42-43). However, the language of RTE did little to stipulate how children's learning should be measured. RTE did not provide clarity on how to operationalize assessing children's learning, neither validating nor invalidating ASER's approach to assessing children's learning with basic metrics of reading and arithmetic proficiency. Figures 1 and 2 are examples of the one-page tests that ASER uses to assess children's learning levels for language and math.[32] The four tasks on each test represent up to a 2nd or 3rd grade level of proficiency, depending on the particular state's curriculum. ASER Centre claims that because RTE fails to offer concrete suggestions for how to assess learning, the Act perpetuates a focus on monitoring inputs into the system versus outcomes of the system in an evaluation of the school system's efficacy.

The main features of ASER's evaluation design along with the rationale for each feature are highlighted in Appendix B. Though ASER reading and arithmetic assessments have not substantively changed since the report's inception, ASER Centre has experimented in some years with additional tests (e.g., comprehension tests and tests of mother's education levels) and survey indicators (see ASER Centre, 2015, pp. 66-67). More recently, ASER has added an English language test (see Figure 11). Furthermore, as the Centre has gained experience in conducting

[32] Figure 1 shows a sample of the reading test in English translation, but in reality, the test would be administered in the regional language of the child.

ASER, its processes for data rechecks and monitoring of data collection in the field have matured (see ASER Centre, 2014a, p. 5). Several of these design features of ASER will be discussed further in the next section on the ASER process.

The ASER Process: Overview

The leaders of ASER Centre frequently boast that ASER lasts only 100 days from beginning to end, of which they are justifiably proud given the enormous labor involved in ASER's execution and publication, year after year. The leadership points out that this short timetable is an important feature of ASER in that its findings, because they are derived from current data, can inform decision-making in nearly real time. Nonetheless, this 100-day estimation pertains to a condensed period from "field to report" starting with the district-level training of volunteers and ending with the national release of ASER (W. Wadhwa, personal interview, May 15, 2015). This short period encompasses a range of activities—training sessions, travel to villages, surveying and testing, monitoring, data recheck and aggregation, data analysis, and report writing and printing—and, it involves the coordination of over 25,000 volunteers and 500 partner organizations along with hundreds of ASER Centre's own staff. Moreover, these activities span the entirety of rural India, so the fact that they are packed into 100 days is truly incredible. Yet, ASER in reality is now a "365-day cycle" that requires the year-round labor and support of ASER Centre national and state teams as the report's underlying process has evolved into its current state of sophistication (W. Wadhwa, personal interview, May 15, 2015).

Before delving into the phases of the ASER process, it is crucial to understand the personnel hierarchy of ASER Centre (i.e., the "ASER Survey Team Structure") that relates to the tiered model of the ASER process. Figure 3 illustrates that ASER is directed from the top by the Delhi-based ASER Centre central team. The central team consists partly of ASER Research

Associates (approximately 15 people) that oversee state teams who are assigned to each of them. Research Associates travel back and forth to their assigned states to help orchestrate logistics and communications between states and the national office (including finance, media, and statistics units) during the ASER process. Meanwhile, ASER state team members are located year-round in state offices across the country (totaling about 64 people nation-wide) and liaison with the leadership of partner organizations and with state government officials. State team members also supervise everyone who executes ASER *within* their respective states, which includes master trainers (MTs) and volunteers (sometimes also called "surveyors"). Across India, roughly 1,000 master trainers are hired each year to facilitate ASER. At the next step down in the personnel hierarchy, MTs coordinate logistics for the ASER evaluation at the district-level, including leading training workshops with help from partner organizations. MTs also monitor the volunteers during field data collection and conduct data quality rechecks for their districts. Meanwhile, volunteers (usually 30-60 per district and approximately 25,000 nationwide) conduct the mapping, surveying, and testing in sample villages, households, and government schools.

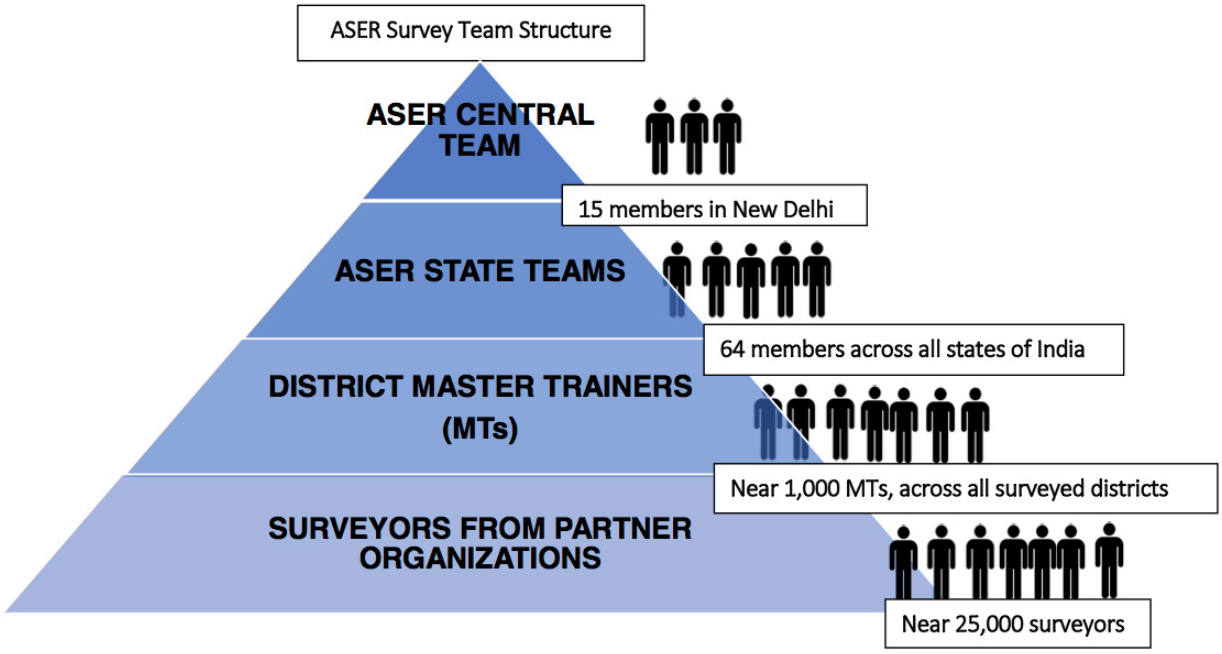

Figure 3. ASER Survey Team Structure. From "Quality Control Framework," by ASER Centre, 2014, retrieved from http://img.asercentre.org/docs/Aser%20survey/Ensuring%20data%20quality/qualitycontrolframework.pdf. Copyright 2014 by ASER Centre. Reprinted with permission.

The Phases of 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										

Figure 4. The phases of ASER process. 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). Adapted from "Quality Control Framework," by ASER Centre, 2014, p. 6. Copyright 2014 by ASER Centre. Adapted with permission.

ASER's annual activities begin in June with forming partnerships and extend into March with the dissemination of ASER findings to various stakeholders at the national, state, and district levels (see Figure 4). Before going to the field, ASER requires an enormous amount of planning. For instance, there is the coordination required to host the National Training Workshop in August, which orients new ASER Centre central and state team employees to the entire ASER process and informs ASER Centre staff (new and old) of changes to tests, forms, protocols, and so on. The following seven sections correspond with the phases depicted in Figure 4, offering details about the extensive process comprising ASER. This study will focus in depth on Phases I-IV, which necessitate the heaviest participation of ASER volunteers, partners, master trainers as well as ASER state and central staff members.

Forming Partnerships and Recruiting (Phase I)

In June, ASER Centre staff members based in all states travel across their respective districts to find viable partner organizations that are willing and able to carry out ASER at the district level.[33] In the upcoming months, ASER Centre needs a partner for every rural district in each of the 29 Indian states (roughly 560 partners in total).[34] Over the years, ASER Centre has partnered with over 2,000 organizations; in ASER 2014, approximately 40% of partners had participated in ASER "at least 3 times before" (ASER Centre, 2014a, p. 6). Generally, ASER

[33] The sampled villages, households, and schools within districts in each state shift over time. For a more complete explanation of how ASER Centre samples households and villages in each state, please reference the NGO's website under "Sampling:" http://www.asercentre.org/Overview/Basic/Pack/History/etc/p/56.html.

[34] Over the course of 10 consecutive ASERs from 2005-2014, ASER Centre has had issues in certain years with securing or maintaining partnerships for some districts. This challenge has resulted in different strategies and outcomes: a) partnering with individuals who are community leaders or very small community-based organizations that operate at the district-level who can help coordinate ASER training and data collection, but who rely heavily on the ASER Centre state teams to execute ASER, b) dropping the district from that year's ASER because of a failed partnership, or c) engaging a partner organization in a neighboring district to complete ASER. Contrastingly, ASER Centre has enjoyed robust partnerships with organizations that span many districts in a state. For instance, Punjab Technical University and its affiliated colleges helped collect data for "14 out of 19 surveyed districts" in Punjab for ASER 2013 and ASER 2014 (ASER Centre, 2014d, p. 6).

Centre forms partnerships with three types of organizations: universities, non-governmental organizations, and government-run teacher training institutes (see Appendix C for the ASER 2014 Partnership Agreement).

Universities and NGOs

Universities and NGOs present some similarities and differences in how they are evaluated and selected as partners for districts. To be selected as a viable partner, both NGOs and universities have indicated to the ASER state teams that they have the experience and organizational know-how to execute something on the scale of ASER at the district-level and within ASER Centre's parameters for quality. In university partnerships, ASER state teams usually forge an agreement with a particular school or unit within the university—ranging from a department of social work to a school of engineering. Typically, the university students of the school or department will serve as ASER volunteers within the district. The students will get trained and ultimately collect the data from households and government primary schools in each of the sampled villages. Universities usually have ample facilities for training but may not have the resources or networks to easily facilitate students' safe travel within the district for data collection. The importance of these factors in forging and maintaining an ASER partnership are highly dependent on the features of the district: for instance, the district's size, its terrain, its general security, and degree of accessibility by private vehicle and public transit can all be vital features in the ease and feasibility of ASER data collection. Partnerships also hinge on the capacities of that state's ASER Centre state team to provide additional assistance and guidance to less-capable organizations if needed.

In the case of NGOs, ASER partners are commonly organizations with some educational or social welfare mission (e.g., human rights, girls' social development, rural livelihoods). These

NGOs usually have experience with fieldwork in villages or have undertaken sizable initiatives in the past. Particularly, critical for the selection of NGOs is whether the organizations have the capacity to conduct ASER given their personnel and resources. For example, one issue is, does the NGO have sufficient staff to execute ASER data collection in 30 villages spanning the district? If not, will the NGO's leadership practicably be able to recruit enough qualified volunteers from their district-area networks or contacts?[35] Another issue is whether the NGO has the facilities to host volunteer trainings. Still, another consideration is if the NGO has the relationships or resources within the district to help facilitate volunteers traveling to remote villages and possibly even staying overnight in them.

Finally, the motivations and preferences of the university or NGO also matter. One major consideration for partners is, do the demands of their regular curriculum or programming realistically accommodate ASER participation? Organizational preferences, motivations for partnership, and regularly scheduled activities can be significant factors in whether the ASER partnership is ultimately viable. Universities must be able to accommodate ASER activities within their program's coursework, testing schedules, and holiday breaks. NGOs must be able to postpone their other programs and services or balance them with the intensity of ASER training and data collection without adversely affecting their beneficiaries and displeasing their funders. The likelihood of getting "buy-in," characterized by strong enthusiasm for and dedication to completing the entire ASER process with quality, from would-be ASER participants—whether those participants are faculty and students in the case of universities, or staff, community

[35] "Qualified" to conduct ASER as a volunteer is initially determined by whether the volunteer is physically able to travel, walk a village, and conduct the surveying and assessment components of data collection and whether a volunteer has the basic education to accurately test children in reading and arithmetic. (In some of the years ASER was conducted, like ASER 2014, a test of basic reading in English was also included so that would be an additional requirement for volunteers.) Ideally, every district has 60 qualified volunteers to conduct ASER—two for each sampled village. Nonetheless, partner organizations can make due with less by conducting data collection over a longer period of time and sending high-functioning volunteer teams to multiple villages instead of just one or two.

stakeholders, and beneficiaries in the case of NGOs—is also a substantial concern in forming the ASER partnership. Such factors are considerations to some extent in forging partnerships with all organizations but are ultimately less pressing in securing the widespread involvement of teacher training institutes, which are unique in terms of ASER partnership and engagement.

DIETs

All Indian states have District Institutes of Education and Training (DIETs) that are charged with teaching and training individuals to become successful primary school teachers in government schools. DIETs are responsible for preparing pre-service teachers and supporting in-service teachers and thus, have a teaching, monitoring, and long-term professional development function in the government school system in their districts. As essentially colleges of primary teacher education, DIETs as ASER partners share some similarities with universities since they help conduct ASER with volunteers that are their own students (i.e., pre-service teachers). DIETs, like universities, generally have the physical space to host the ASER training and the organizational capacity to coordinate the data collection.

What currently makes DIET partners unique amongst ASER partnerships is their engagement is not typically voluntary and their students' status as future teachers.[36] With few exceptions, DIET partnerships are now negotiated at the state-level with ASER Centre state teams through the State Council of Educational Research and Training (SCERT). SCERTs control all DIETs within their state. After ASER state teams form a partnership with their respective SCERT, the SCERT, in turn, sends letters to all of its DIETs that mandate their

[36] The involuntary nature of DIET partnership in ASER is relatively new. Before ASER 2012, DIETs who participated as partners in ASER did so because they were approached by state team members and consequently, opted to join in ASER efforts. In those earlier ASER partnerships with DIETs, the above-mentioned factors (e.g., organizational capacity, resources, and interest) that were influential in how one evaluated the viability of a partnership were presumably more central to decision-making on both the DIET's part and ASER Centre's part in whether the partnership was ultimately formed.

participation in ASER training and data collection for that year. For ASER 2014, "243 DIETs across 12 states" served as partner organizations, resulting in the involvement of nearly "14,000 future teachers" (ASER Centre, 2015, p. 63). According to ASER Centre, this partnership is highly beneficial for both ASER Centre and DIETs because DIET students 1) acquire "firsthand understanding of children's learning levels in their own districts" and 2) learn the "simple ASER method of assessing learning levels of children which they could utilize as future [government] school teachers" (ASER Centre, 2014a, p. 7), including techniques of "building a child-friendly environment before testing" (ASER Centre, 2015, p. 63).

While ASER Centre state teams prefer to have partners identified and partnership agreements wrapped up by early August, it is not possible for some states in some years. For instance, Tamil Nadu's state team was still in the process of identifying partner NGOs or universities for some districts well after mid-August. In a large state such as Tamil Nadu, identifying partners for all of its 32 districts is a considerable feat since (as of ASER 2014) the Tamil Nadu SCERT has not mandated DIET participation.[37] Finding partnerships in Tamil Nadu requires extensive research into what types of organizations exist in each district, and ultimately it entails significant travel for the ASER state team within districts to meet with potential partners, evaluate their abilities and interest in executing ASER, and negotiate agreements and initial plans for how and when training and data collection will be undertaken.[38]

[37] In fact, the state leadership of Tamil Nadu has been typically resistant, if not hostile, to ASER results because the results have often indicated learning levels much lower than what is widely anticipated in one of the purportedly best states for education in India. It is popularly believed that Tamil Nadu's more socialist-leaning government is very strong in the provision of publicly funded programming and systems, and thus ASER data is often challenged by government officials in that it undermines a generally favorable view of government education and public administration in the state.

[38] Ultimately, ASER Centre's level of selectivity with partners can depend on the number of potential partner organizations in the district and their degree of interest in ASER participation. Compromises may be made in the level of capacity or experience of a partner organization or in its number of readily available, qualified volunteers in the interest of forming some type of partnership and having some assistance with ASER in the district. The ASER Centre state teams make these calculations with guidance from the ASER Centre national team.

Recruiting Master Trainers

Recruiting master trainers is an equally important aspect of this first phase of ASER (see Appendix D for the "Roles and Responsibilities of Master Trainers"). MTs, as they are referred to throughout the ASER organization, come to be involved in ASER through typically four different pathways: 1) they are affiliated with the district's partner organization for ASER; 2) they are independently recruited from universities, via advertisements, or through some other mechanism (e.g., word of mouth); 3) they are an employee, volunteer, or former employee or volunteer of Pratham; or 4) they are a returning MT from past ASERs. ASER Centre generally tries to hire two MTs per district, but the ASER state teams sometimes have difficulty finding and retaining interested and qualified candidates. MTs are employed for only a few months during the ASER season and are paid a relatively low stipend for the amount of travel and the workload required. Moreover, MTs need to have a formal education, demonstrate some level of comfort with data and research protocols, and possess organizational and supervisory skills.

Fidelity to ASER protocols and obtaining correct data at the district-level is the primary responsibility of MTs; it is their topmost objective as it relates to the ASER process and their interaction with partners and volunteers. To ensure MTs' unfailing commitment to this objective, ASER state teams have employed different strategies in their recruitment. In some states, ASER state teams prefer that MTs do not come from the partner organization itself, especially if the district's volunteers are from the partner organization. The logic is that external MTs provide an extra layer of oversight and assurance in the quality of the ASER process, but instead, if MTs share an affiliation with both the partner organization and the volunteers for their district, MTs may have conflicts of interest and be less likely to scrutinize the quality of data or actively correct sloppy work or shortcuts in the training or operationalization of ASER. In other states, ASER state teams have experimented with alternative approaches, such as a hybrid recruitment

strategy for each of the state's district, comprising of one "local Master Trainer recruited from the partner organization of the district" and one "external Master Trainer" who is from a "central partner" and "allocated districts to measure across the state" (ASER Centre, 2014a, p. 9). Regardless of strategy, ideally all partnerships for districts are formed and all MTs and volunteers are recruited (independent of the partner organization, as needed) prior to beginning the phase of training workshops for ASER.

Training to Conduct ASER (Phase II)

Every year, ASER includes a three-stage, three-tier training process that starts with the ASER National Workshop, then proceeds to state-level training workshops, and concludes with district-level training workshops (see Figure 5). All training sessions are aimed at helping people at each level of ASER (national, state, and district) perform their roles and responsibilities for ASER's sound implementation. Figure 3 depicts the structure of ASER Centre personnel, partners, and surveyors, who are involved in the ASER process at different tiers, and Figure 5 conveys how they are all trained in three separate phases in order to conduct ASER. The ASER Centre national (i.e., central) team members host the National Workshop for training the respective state team members. Then, the ASER state teams each host one training workshop in their respective states to prepare their master trainers. After state-level training workshops, every state holds numerous district-level training workshops to train volunteers (i.e., surveyors); district-level trainings are facilitated by the master trainers assigned to that district with the support of partner organizations.

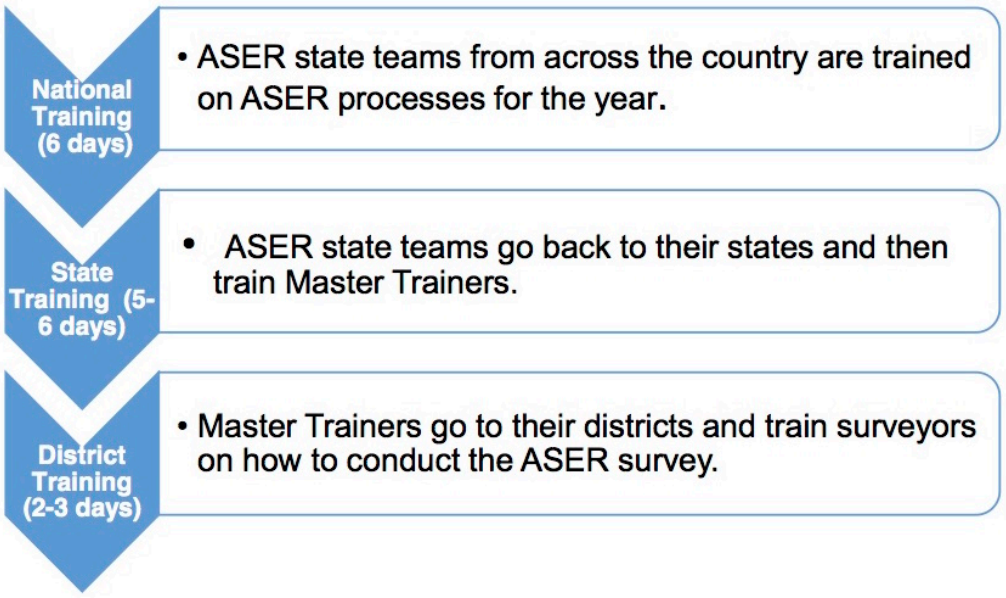

Figure 5. ASER three-tier training process. From "Quality Control Framework," by ASER Centre, 2014, retrieved from http://img.asercentre.org/docs/Aser%20survey/Ensuring%20data%20quality/qualitycontrolframework.pdf. Copyright 2014 by ASER Centre. Reprinted with permission.

National Workshop

ASER Centre typically holds its sizeable, six-day National Training Workshop in August, marking the beginning of that year's cycle of training, data collection, and monitoring. The purpose of the workshop is to cover all aspects of the ASER process for ASER central staff and state teams in order to ensure timeliness and quality in that year's data collection for the annual report. Nearly everyone who has a role in the ASER process at the central and state levels attends the workshop. The following two sections give further details on the National Workshop: 1) its participants and training environment and 2) the content and instructional approach of its training sessions.

Training environment and participants. Staff members from across India as well as guests—e.g., Indian bureaucrats, foreign researchers and university-student interns, and representatives from foundations and think tanks—attend the workshop facilitated by ASER Centre. In 2014, the workshop attendees totaled roughly 110 staff members and 25 guests. The

ASER National Workshop is akin to a giant family reunion with people from all corners of vast India coming together after a year apart to greet one another, exchange stories, introduce new members of state and central teams, and tease each other about their cultural differences from food preferences to language. Indians (and foreigners), who would otherwise never meet, share living quarters and meals and coach one another through the parade of training sessions and field visits (to pilot the ASER data collection protocols in villages). The workshop usually makes a profound impression on first-time attendees, such as Ruchi,[39] a newly hired research associate in the Delhi office. In her ASER blog post entitled "All Serious Daring Starts from Within," she states:

> I am Ruchi and I work at ASER Centre, Delhi. This will be my first ASER."[40] Yes, exactly these lines marked the beginning of the ASER-2014 National Workshop for me. Little did I know what the days to follow will hold for me but I was sure of one thing- the unexpected was about to turn real. Of long hours and late night mock training preparations. From covering hamlets kilometres apart during pilot to getting drenched in the rain during field recheck;[41] the workshop brought with it experiences that I could go on talking for hours about but I will still fall short of what it exactly meant to me. So much to say yet so little words to express. And now we are back in Delhi and I am all unpacked. It is 2:18 am (IST) and I found myself smiling thinking of all the wonderful experiences that I had in Aurangabad and all the memories I brought back with me.[42] I am glad to be part of the making of the 10th year of the ANNUAL STATUS OF EDUCATION REPORT. As we prepare our timelines and work on our budgets and gear up for a season of running around– working our way through all the [data collection] packs and deadlines- I wish each one of you reading this (or not) all the very best for a grand and successful season of the survey. As for me, I am excited at the prospect of being part of the making of history and nervous because that just means the stakes are really high. But then we are all together in this and I am confident we can – "DO THIS". See you all on the other side. Good Luck. (August 10, 2014)

[39] The name of the ASER Centre research associate has been changed to a pseudonym here, although the original blog post was public and included the full name of the associate.

[40] Ruchi's introductory quote is pertaining to the welcome sessions of the National Workshop where everyone introduces themselves, the state they represent, and the number of ASERs in which they have participated.

[41] Ruchi is discussing part of the training, which requires mapping all of the hamlets (different sections) of a village, before a sampling of households can be taken for surveying. This will be discussed further below.

[42] IST is India Standard Time, and Aurangabad is a city in the state of Maharashtra where the ASER 2014 National Workshop was held.

Ruchi's blog hints at the intensity of the training sessions with their "long hours and late night…preparations," and she describes the hard labor of field piloting ASER data collection and recheck protocols that require mapping "hamlets kilometeres apart" and "getting drenched in the rain." In addition, Ruchi expresses how the National Workshop seems to be the beginning of many formative life experiences to come that will result from her all-consuming participation in ASER—experiences that she "could go on talking for hours about" but "will still fall short of explaining what it exactly" they mean to her.

It is at the National Workshop that new ASER employees get a realistic and more experiential understanding of what they must accomplish over the next few months of "timelines," "budgets," "deadlines," and "running around." The sheer magnitude of work ahead is quite exhilarating and nerve-wracking, as is the fact that for some ASER staff (in their early 20s), this is one of their first forays into the professional world after college or a masters program, and their first experience doing rural fieldwork, and their first assignment in a supervisory role. This is what seemingly prompts Ruchi to wish each of her counterparts, "Good luck." With perhaps still a limited sense of what she is heading into, the concluding lines of Ruchi's blog post about "high stakes" and being "all together in this" convey an eagerness to go forth into the world with a commitment to something altruistic, to something momentous, to "the making of history." Her words illuminate a common hope often embedded in the work of ASER—from its central staff to its volunteers. Ruchi's belief in the power of civic engagement shapes her motivation for joining ASER Centre. Her perception of high stakes reflects both the girth of tasks and responsibilities that are placed on her youthful shoulders, but also the monumental personal journey that awaits her and other ASER central team members, who will see each other "on the other side," after having sacrificed the privileges and comforts of their

typically urban, upper-middle class lives to travel by train, bus, motorbike, and foot through remote terrain and inclement weather to survey villages in a part of their own country that they, for the most part, do not know.

It is important to note that ASER state team members, unlike ASER central team members from Delhi such as Ruchi, often have less affluent or privileged backgrounds (though they are typically still economically and educationally privileged relative to the average person within their state). Also, whereas the turnover of staff who participate in ASER from the central office is quite high, the state teams are usually comprised of many members who have served through several ASERs and are therefore older and more seasoned when it comes to fieldwork and more knowledgeable of rural India generally. State team members also tend to be male whereas the central team has more of a gender balance. Nonetheless, there are exceptions to my characterizations of both central and state team members, and in the case of newly minted state team members, they can have many of the same reactions as Ruchi to their first National Workshop and field experiences with ASER. Furthermore, state team members travel just as much, if not more, during the ASER data collection, monitoring, and recheck processes.

The National Workshop takes place during the end of the hot summer months in an often un-air-conditioned vocational training facility operated by ASER Centre's parent organization, Pratham. These vocational facilities are located in northern or central India, so temperatures range from average highs in the mid-90s F to mid-80s, but depending on the monsoon patterns, temperatures can spike much higher. The workshop accommodations and training space are neither meager nor plush by conventional Indian standards. ASER staff members stay in simple, large dorm-style rooms with rows of bunk beds or mattresses on the floors. Usually, communal bathrooms and showers serve an entire floor. Internet service and steady electricity are typically

wanting. Meals are eaten in a large dining hall where participants bus and sometimes clean their own dishes.[43] Depending on the particular facility for that year's workshop, the workshop welcome, guest keynotes, concluding assembly, tests, and whole-group training sessions are held in a large hall where everyone sits cross-legged on the ground upon large mats and rugs or in plastic chairs assembled in rows. The presentation space has a projector, screen, sound system and podium or table with chairs.

While attending the National Workshop, a much discussed topic amongst ASER staff members is the menu: mostly vegetarian food is served like curried lentils or chickpeas (*daal* or *chaana*), wheat flatbreads (*chapaati, roti,* or *naan*), thin yogurt with onions and tomatoes (*raita*), and stewed vegetables (*subzi*). The menu reflects a simple, but routine diet for many north Indians. However, especially for some state teams, the repeated absence of meat, rice, and more familiar spices is difficult to stomach over six days, stoking feelings of homesickness or causing digestive problems. A special meal (generating outward enthusiasm) is one that features some of these missing items—particularly, meat. Regular tea (*chai*) breaks with cookies (*biscuits*) and meals that end with sweets are small but appreciated comforts.

During the upcoming data collection period, ASER Centre staff members will all prove their resilience and hardiness, as they have time-and-again, by managing the harsh travel and heavy workloads required. The climate, food, and living conditions at the National Workshop are the onset of that conditioning for some: it is often hardest on two groups—new central team members who are unaccustomed to this type of travel and fieldwork, or state team members who traveled long distances on the train from northeast and southern India. The latter state team

[43] For example, the national training for ASER 2013 was held at the Pace Hospitality Training Center in Siwan, Bihar, which is essentially a vocational school where high-achieving young people from historically marginalized and minority communities (recruited from both rural areas and urban slums) learn the skills required to hold jobs in the service and hospitality industries, such as hotels. Thus, students in-training at the Pace Center were cooking the food and providing the housecleaning for the National Workshop.

members arrive frequently exhausted and must adjust to food, water, and weather different from what they are accustomed. Usually, a handful of National Workshop participants get sick—exhaustion, dehydration, and dysentery seem to be the most common complaints, though more extreme cases of intestinal infections, flu, and other ailments have occasionally sent ASER staff to hospitals in nearby cities.

Content and instructional approach of training sessions. The ASER National Workshop's format and curriculum do not substantially change from year to year, but rather, are in a gradual process of refinement. The training days are divided into four days of classroom-based training and two days of field-based training. The field "pilot" visits offer participants the opportunity to practice village mapping and surveying, school surveying, and household surveying, and testing of children (within the strict parameters of ASER Centre's protocols for each of those methods.) The instructional approaches employed in the training sessions range between lecture, modeling (or role-playing), collaborative learning (or group work), and experiential learning (in the field). Built into the conference schedule—or sometimes, urgently carved out of the schedule to the detriment of participants' sleep—is time for practicing how to train others in the ASER process (i.e., mock training) and also time for learning how to document the ASER process (via blogging, social media, photography, and videos).

Figure 6 displays the workshop topics by day. The workshop's execution necessitates the efforts of people across the ASER organization and its ultimate success relies upon intensive teaching and learning activities aimed at delivering copious amounts of content. Members of the ASER central team in Delhi who have experience in conducting ASER over at least a couple years' time, generally coordinate the training logistics, determine the session schedule, and divide national and state team staff members into the Hindi and English language groups for the

workshop.[44] ASER Centre directors and distinguished guests open the workshop and deliver welcome addresses that are intended to motivate ASER staff members—both new and old—by highlighting the importance of ASER to the nation's education system.

Many sessions are facilitated by ASER state team veterans who, over successive years of conducting ASER, have distinguished themselves as experts and unique talents in certain areas of ASER training and implementation. For instance, one veteran is golden-tongued when it comes to communicating with the Sarpanch (i.e., village head) or school headmaster—these are key stakeholders, whose cooperation is necessary for ASER surveying—and therefore, that ASER veteran leads the session on how to explain what ASER is to important gatekeepers within villages. Alternatively, another ASER veteran is particularly adept at teaching how to map villages and how to systematically select far-spread households on mountainous terrain; thus, this ASER veteran helps lead the training session on the household mapping and sampling process. These are amongst the most important sessions describing ASER's "nuts and bolts," which directly affect the quality of its data.

Generally, ASER training sessions are jam-packed with information—the rules, the paperwork, the data collection formats, and the veteran's tips for dealing with all extenuating circumstances—aimed at ensuring uniformity and quality in the execution and data collection of ASER in extremely different state environs. How to cope with specific challenges that pertain to travel, gaining access to villages, sampling, or surveying and testing are discussed at the national

[44] India is an extremely multilingual country and ASER Centre National Workshop training sessions must contend with the complexity of bridging this linguistic pluralism. ASER Centre national staff members based in Delhi are usually proficient in both Hindi and English—many national staff members even speak regional languages like Tamil, Kannada, Punjabi or Bengali as third or fourth languages. However, ASER Centre state teams members are frequently most comfortable in the language that is dominant within their state (e.g., Gujurati, Manipuri, Oriya, and Hindi) and then, if their state language is not Hindi itself, they are either proficient in Hindi or English as these are the two official national languages of India. Nonetheless, proficiency levels and degrees of comfort with Hindi or English does vary greatly amongst staff. One of the advantages of ASER Centre employment for some state team members are the linguistic capacity-building opportunities that training and fieldwork provide.

training—usually through people exchanging stories and passing on pieces of wisdom tied to difficult scenarios. However, obstacles that are particular to a state or handful of districts are usually discussed later and in-depth by state teams and partner organizations at either the state-level or district-level training workshops.

ASER National Workshop Sample Schedule of Sessions

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
Session Type & Location	*Classroom-based sessions*	*Field pilot survey & testing*	*Classroom-based sessions*	*Quizzes & field data recheck*	*Mock Training sessions*	*Classroom-based sessions*
Training Topics & Activities	Welcome	Field visit preparation	Field Visit Feedback	Monitoring & Recheck quiz	Reporting data from field pilot surveys & tests	Monitoring & recheck quiz results, review & clarifications
	Introduction to ASER	Travel to village	ASER Data collection quiz	ASER data collection quiz results		
	Meeting the Sarpanch	(Pilot) Meet the Sarpanch	Monitoring processes	ASER data collection theory, review, & clarifications	Mock Training: Meeting Sarpanch	ASER sampling
	Village map & survey	(Pilot) Village mapping & surveying	Recheck processes		Mock Training: Village Map & Survey	Closing remarks
	Household survey	(Pilot) Household surveying	Financial management	Field Visit Preparation	Mock Training: Household survey	
	Reading & math learning tests	(Pilot) Testing children		(Pilot) Field Rechecking data	Mock Training: Learning Tests	
	School survey	(Pilot) School surveying	Documenting ASER	Compiling recheck field notes		
	Data entry in test & survey Forms	Compiling field notes & data in forms		Mock training preparation	Mock Training: School Survey	

Figure 6. ASER National Workshop sample schedule.

Sessions on each day of the National Workshop are scheduled from roughly 8:30 a.m. to 8:30 p.m., but they frequently last (longer) into the night. The daily schedule is easily derailed in the course of addressing questions and offering clarifications during sessions that are dense with crucial details. Rules and processes must be understood thoroughly by everyone and committed to memory before the state team members leave the National Workshop, as they, become the

ASER trainers for MTs and thus, key disseminators of vital information about ASER protocols. During the course of my observations at three National Workshops (for ASERs 2013, 2014, and 2016), I have seen sessions last well past 10:00 p.m., with a short break for dinner, and then to resume for hours afterward. Moreover, after formal sessions wrap-up, other meetings begin between small groups or partners that are planning, for instance, mock training sessions for the next day. On some days of the Workshop, participants are not getting much sleep as they are also trying to manage logistics within their states for the upcoming ASER state-level training workshops.

On the day prior to the start of the National Workshop, ASER state team members arrive by trains and hired vans to the training location. That night of Day 0 is for settling in and catching up with one another as the central team from Delhi rushes around finalizing plans in the hours before the Workshop sessions begin. Day 1 of sessions covers all the theory and protocols underlying the ASER evaluation design and its process for collecting data in villages, households, and schools. Workshop participants learn how to explain what ASER is, how to map the village and systematically sample households, approach households for surveying, and then conduct the survey. The curriculum for the second half of Day 1 details how to interact with children, make them comfortable during testing, and administer and score the reading and arithmetic tests consistently and accurately. Afterward, state team members learn how to conduct the ASER survey in government schools, noting what data they take from school registers, from observations, and from the mouth of the headmaster or teachers. Finally, before Day 1 wraps, participants are trained on how to fill out the test and survey forms properly and aggregate data for villages and then for districts.

On the very next morning, Day 2, everyone travels in small teams to neighboring villages (generally within one to three hours' distance) to pilot the entire mapping and data collection process that they learned just the day before. Usually, two or three experienced ASER team members lead the small pack of trainees through the process once in the village. For new ASER team members, this can be their first experience in a village and for many ASER participants—new and old—it is their first time in a village in *this* part of rural India.[45] Thus, the pilot is often a cross-cultural experience, which can be exhausting, uncomfortable, and perhaps, anxiety provoking, but also, thrilling, heartwarming, and mind-expanding. In a blog post entitled "The Simplicity of Impact," Hema, a former teacher and new ASER research associate from Delhi, explains how her experience with the pilot field visit challenged her ideas: "This is a story…for those who are often narrowed by their perception of villages and its people, including myself, but now feel slightly more aware and enlightened." She describes how in a short time her ideas evolved:

> It's been a month since I became a part of ASER Centre. As a believer of organized systems, laborious methodology and stories of impact, I could not understand how a simple floor-level text such as the one ASER employs, could create tangible impact. But now I do…. I recently had a hands-on experience with the simplicity of impact. It was quarter to 10 in the morning; and I was all set to go with a team to a nearby village in a block in Aurangabad district to conduct the ASER [pilot] survey. I was ready with the testing tools, eager, scared, hoping to diligently execute each and every part of the process. Going through a meandering course, we finally reached our designated destination, a cute little village nestled in the Western Ghats. Chai, gupshup,[46] smiles and some Marathi translations later, we find ourselves at the first household to survey. We were welcomed by the grandmother with a warm smile and a twinkle in her eye. We gave

[45] Part of the difficulty in orchestrating the logistics of the field pilot visit (and a sign of the conscientious planning done by the Centre's staff, year after year) is that language differences must be bridged when practicing data collection in villages. Thus, state team members from outside the "Hindi belt" are often relying on their Hindi-speaking counterparts to help facilitate their interactions with villagers during the pilot because they generally take place in a Hindi-speaking state. In 2014, however, when the National Workshop was conducted in the state of Maharashtra, the central team had to arrange for extra volunteers and employees from Pratham to help facilitate the field visits as few people were able to speak the local language, Marathi, which was the mother tongue of parents and children in the pilot villages.

[46] Hindi for *gossip*.

> the kid of the household a subtraction problem while the grandmother sat behind her and watched with eager eyes. She sat with bated breath as she watched the kid move up to the division problem. When the kid finally completed the test, the proud grandmother shared her aspirations for her granddaughter. As we continued the conversation, with the help of broken Marathi and gestures, I got to know a little more about the family. We chatted about the importance of attending school every day and the impact it can have on the child's life. The girl's parents were both educated, I was told, and I could sense the pride in their voices while talking about their family and how everyone in the village was in awe of this fact. As I walked away, it struck me that a simple testing tool had facilitated a meaningful conversation about the importance of education. The ASER testing tool reaches the hands of more than 6 lakh children every year.[47] Think about it. (August 29, 2014)

The account Hema offers of her pilot experience conveys that while Workshop participants are learning the technical aspects of how to carry out ASER data collection properly, they are also having intimate interactions with families in their homes. Testing children in households, in front of their families, as the ASER design requires, can stimulate heartfelt conversations about children's learning, "the importance of attending school everyday," and a family's "aspirations for their granddaughter" even through differences in language and background. Hema's experience of connecting with the village's families through "broken Marathi and gestures" on the "importance of education" altered her ideas about what "impact" means and her notions about who rural communities are.

During the pilot, National Workshop participants complete all phases of data collection that they will fulfill during the real ASER; however, the pilot is completed in one day versus two days because there are generally triple the amount of people to complete mapping, surveying and testing (i.e., six participants during the pilot versus two ASER volunteers during the actual ASER data collection). When the small field teams return to the training center—usually after dark, they will huddle together to aggregate all of their household, testing, and school survey

[47] A lakh is common numerical unit in India, which equals 100,000. Thus, six lakh children is 600,000 children.

data into one set of forms so it can be reported to the Workshop coordinators and utilized on Day 4 for data recheck. The ASER Centre directors typically also request the small teams to produce a written reflection and submit photographs from their day of surveying and testing, which can be posted to the ASER blog and used in that year's final report or media materials.

Day 3 of the National Workshop begins with a test on all the protocols, forms, and rules comprising the ASER process that have been taught thus far. Afterward, sessions are held on how to properly monitor volunteers' and MTs' work as well as how to train MTs to identify and then prioritize weaker volunteers, who need extra monitoring and support. The ASER Centre directors encourage state team members to pair weaker MTs with stronger MTs when deciding which MTs will be partners and in assigning them to districts. State team members are advised to give less difficult district assignments (e.g., districts with less remote villages and more experienced and competent partner organizations) to weaker MT pairs, so they are likely to confront less challenges during data collection and can be more easily monitored and supported by state team members. (This same strategy is taught later to MTs during the state-level training. State teams members in turn, will advise MTs on how they pair up weaker and stronger volunteers and assign volunteer partners to surveying sample villages within their district.) Next, there is a session on how to conduct data rechecks in-person, by phone, and in villages, and how to decide when it is necessary to re-survey a village. Finally, there are sessions on the management of finances (e.g., budgeting, paying MTs stipends, and reimbursing partners' and volunteers' costs) and on documenting the ASER process effectively (e.g., how to take compelling photographs and videos of data collection).

The major event of Day 4 is returning to the field for piloting how to do data rechecks in villages. Small groups are allocated another group's village and pack of data forms (from Day 2)

with the assignment of rechecking their school and household survey data and confirming that children were tested according to protocol (i.e., in their home, following the proper question sequence for test sections). Before going to the field in the afternoon, there is also a quiz on the monitoring and recheck process and a financial management quiz. Following the quizzes, there is a session that reviews the correct answers (and rationale) for the first quiz given on Day 3 regarding how to collect ASER data. In the evening, after participants return from the pilot data recheck in villages, they pair off into their teams for mock training preparations.

Days 5 of the National Workshop is mostly dedicated to giving participants the opportunity to practice teaching ASER modules. This practice is necessary ahead of state-level training workshops. Those experienced in ASER—from ASER Centre directors to ASER veterans on state teams—serve as coaches (and sometimes critics) of the "mock trainers," who are judged on many aspects from their style of delivery and ability to keep their audience's attention to the clarity of their explanations and the comprehensiveness and accuracy of the information that they convey. Mock training sessions serve several purposes: 1) they build inexperienced trainers' courage to speak in front of a group and deliver content with confidence; 2) they provide an opportunity for some state staff to practice presenting in a second or third language;[48] 3) they reinforce knowledge of content by requiring participants to have enough understanding to explain and teach it to others; and 4) they present another opportunity for experienced ASER staff members to clarify processes and to share tips with their more junior counterparts.

[48] Mock training sessions are conducted in either Hindi or English. However, there are many states in which the main language is neither of these two. For instance, in the northern Indian state of Punjab, the main language is Punjabi and Hindi would be a second language for most people (if spoken at all). Similarly, in the southern state of Tamil Nadu, Tamil is the main language and English is a second language (if spoken at all) or perhaps, third language behind Telugu or Malayalam. Though state teams will complete mock training sessions at the National Workshop in either Hindi or English, they will present ASER content in their state-level training workshops in whatever language is most appropriate to their state.

Day 6 of the National Workshop is dedicated to tying up loose ends, so to speak. There is a review of answers from the monitoring and recheck quiz with clarifications provided. Also, there is a session that explains in more depth ASER's sampling approach (Probability Proportional to Size or PPS).[49] Next, a session is held that discusses ASER cases studies from the field, which illustrate further points regarding its process. Finally, the National Workshop ends with the ASER pledge, which states (in part):

> I pledge to carry forward the honest work and vision of those who have come before me and become an example for those who will follow me in the years to come…. I pledge to uphold the integrity of the important role this movement has entrusted in me and play my part in building a better India. (see Appendix E)

The pledge emphasizes personal integrity and ASER's larger, noble purpose—as "this movement…in building a better India"—it is a purpose that staff members are directly charged with fulfilling. In the pledge's text, individual accountability is linked to the quality of the data and fidelity to the correct protocols: "I understand the importance of recording only correct information in the survey sheets and pledge to do so with complete honesty." One reason why the pledge emphasizes honesty and doing work with "the utmost sincerity" is that the demands of the ASER process can be extreme, but if failure to adhere to the protocols frequently wins over devotion to the prescribed procedures for ASER (no matter the circumstances or obstacles encountered), the resulting data will not be accurate (i.e., reliable and valid across the over 16,000 villages surveyed annually). Thus, the Workshop ends with all ASER staff vowing that "under no circumstances, will I avoid my responsibilities during the course of the survey" as they will in turn ask their master trainers and volunteers to vow over the coming weeks in the state-

[49] For more information on ASER's sampling approach, see ASER reports or the paper, ASER Centre (undated.) "Sample Design of Rural ASER," available at http://img.asercentre.org/docs/Aser%20survey/Sampling/Sample_Design_of_Rural_ASER_1.pdf

level and district-level training workshops, which will all conclude with the same pledge to ASER.

State-level and District-level Training Workshops

The state-level and district-level training workshops follow the National Workshop in succession. The state-level workshops are held over roughly a five-day period, and the week or two between it and the National Workshop is an extremely hectic time for state teams. An ASER employee from the state of Assam, Sujatha,[50] describes her monsoon-related challenges in planning their 2014 state-level training workshop in a blog post entitled, "The Excitement Never Stops…"

> I would like to describe one memorable day in the [ASER Centre Assam State] office. Just four days before the state-level training in Assam, my team members and I were busy preparing for the five hectic days of training ahead. I was on the phone inviting the district DIET principal to attend the workshop. But my efforts were falling on deaf ears. The principal was busy with other projects that coincided with our event. So attending a five-day residential training was out of the question. I was disheartened. So, we decided to issue a fresh letter of invitation to all the DIET principals in the state. By this time, heavy rains had begun and within minutes, the office premises began to get flooded. My colleague, Deepali, and I began to panic. We had to make our way back home quickly. At about 5 pm, we made our way out of office by hiring a rickshaw where we had to sit on top of the rickshaw to avoid being submerged in the rising flood. Upon returning home, I watched flood water enter my house and my eight-year-old, who was trapped inside, calling out to me for help. Just then, the phone rang and it was the DIET principal. Cradling the phone between my ears and shoulder, my files and documents strapped between my arms, I made my way into the house and lifted my son out of the water, all the while just happy to hear the principal confirm his attendance at the state level training. It was all worth it. This is the 10th year of conducting ASER in Assam, and the excitement never stops. (August 30, 2014)

A quality that frequently defines the work of ASER state team members during the training, data collection, and recheck phases (II-IV) is on display in Sujatha's above blog post—this quality is a relentless, selfless dedication to achieving the objectives set for one's work. Sujatha's description that emphasized above all else her relief that the head of their partner organization,

[50] Sujatha and Deepali are pseudonyms.

the DIET principal, was going to attend the ASER Assam state training workshop after all, may see a little insane as she stands in water, cradling both her son and her ASER paperwork as her house floods. However, this self-sacrifice and capacity to keep moving forward despite sizeable obstacles, is often necessary to (literally) weather the challenges of climate, political instability, unreliable partners, and so on that are routinely part of ASER. Sujatha's resilience is an essential characteristic of a successful ASER central or state team member.

The state-level training workshop will follow a similar sequence of sessions as the National Workshop with basically the same curriculum and instructional approach (omitting only a few topics that pertain solely to the ASER central or state teams). As mentioned, state-level training workshops are run by ASER state team members and dedicated to training MTs on everything they need to know in order to facilitate district-level training workshops and then supervise volunteers during surveying and testing in villages. MTs need to learn how to monitor data collection and recheck ASER data for their respective districts. Throughout the state-level training, MTs are observed and given tests by the state team to assess their comprehension of ASER protocols and their readiness to assume district-level responsibilities. Once the state-level training workshop is finished and MTs have made their ASER pledge, they are handed their district assignments. As mentioned above, two MTs are assigned to each district and a common strategy employed by ASER state teams is to pair weaker MTs with stronger ones in order to try to ensure quality training, supervision and monitoring in the district. Also, more difficult districts (e.g., more remote, with greater linguistic diversity, and with weaker partners) are generally allocated to a comparatively stronger and more experienced pair of MTs. The gender of MTs can also influences who is assigned to which district.[51]

[51] Issues of gender in ASER participation will be analyzed in future research, but my preliminary analyses of data from my wider study of ASER reveal that in some states, MTs are disproportionately male and where

District-level trainings are held throughout states following the completion of state-level trainings. Often the district-level trainings are held at the facilities of partner organizations. Depending on the type of partner and whether the volunteers come from the partner institution, the atmosphere and participants of a district-level training can vary greatly. For instance, in the Tamil Nadu district of Madurai, the partner organization is People's Watch—a human rights organization. In ASER 2014, the MTs for Madurai district were experienced members of the People's Watch staff and the volunteers were a hodgepodge group of local young people recruited from teacher training universities and other academic programs (e.g., law school) as well as some of the organization's own or affiliated members. However, in states with DIET participation, volunteers are typically pre-service teachers from that partner DIET for the district. Training workshops are held at the DIET itself, all ASER activities are coordinated out of the DIET, and MTs (while often independent from the DIET) interface heavily with DIET principals or coordinators in facilitating the volunteers' participation in the district data collection.

District-level training workshops are generally three days long with one field pilot day for volunteers to practice all methods of data collection. Like the national and state-level workshops, the participants (i.e., volunteers) are observed and assessed to make sure they understand all mapping, surveying and testing procedures in villages. This includes a test that assesses volunteers' own ability to execute and interpret the reading and math calculations required by the ASER tests (Figures 1 and 2). After completing the training sessions and taking

---

women are involved as MTs, their security is frequently a focus in both the assignment of districts and in planning their travel to villages. For instance, in the state of Rajasthan during ASER 2014, there were only two female MTs out of roughly 50 MTs hired for conducting ASER. These two female MTs were assigned the same relatively central and urbanized district where they did not have to routinely travel to remote areas, have to travel throughout the night, or have to stay away from their homes. Another accommodation for female MTs (that some ASER state teams utilized) was a male state team member accompanied female MTs during travel to villages and monitoring. Such decisions appear to be dependent upon many factors including state-specific gender norms and security concerns as well as female MTs' level of comfort with travel.

the ASER pledge, volunteers are assigned to a village from the 30 sampled villages on the confidential district list (generated by ASER Centre headquarters in Delhi). Two volunteers survey a village together. They are paired with one another utilizing the common partnership strategy predicated on one member being stronger and the other weaker. MTs along with the assistance of partner organizations and ASER state team members are in charge of devising a game plan for how to tackle all sample villages given the number of volunteers for the district.[52] MTs and ASER state team members make sure volunteers are appropriately monitored.[53]

Collecting Data and Monitoring of Data Collection (Phase III)

Following successful completion of the district-level training workshop, a pair of volunteers is assigned a village for "survey," i.e., for collecting data for ASER. The two district MTs (or sometimes, partner organization) will help the volunteers figure out where the village is located in the district and how best to reach it—by public or private bus, by motorbike or hired vehicle, on foot, or via some combination of the above. Volunteers will receive a small stipend to cover their transportation costs and food for their two survey days—generally, a Saturday and Sunday.[54]

[52] Recruiting and training 60 volunteers for a district is ideal so that each pair of volunteers only surveys one village. However, it is not always possible to recruit this many and pairs may need to survey multiple villages.

[53] As with MTs in the above footnote, female volunteers are often given closer villages where travel is less precarious and in which it is easy for them to leave from their home in the morning and return well before dark. More remote places are assigned to male volunteers and the most enthusiastic volunteers as the challenges to reaching them can be several—some villages in Northeastern India require extensive hiking to reach them (and may even necessitate having special language skills to communicate with tribal communities). In such situations, volunteers cannot easily leave and return to the village for a second day of survey, requiring that volunteers arrange to stay overnight with someone in the village. In these cases, an ASER state team member or the district's MT may accompany volunteers to complete data collection in that village and stay overnight with them.

[54] The two-day weekend survey period is a general model for ASER that is fluidly applied. In certain states (e.g., Manipur) or districts a different approach is best for feasibly and reliably collecting data.

Meeting the Sarpanch

Once the two volunteers arrive in their assigned village, the data collection process begins. First, they try to locate the Sarpanch (i.e., the elected head of the Panchayat, which is a village-level governing body) and meet with him or her in order to gain support for completing the survey in the village. This is generally accomplished via a sit-down meeting where the volunteers offer an explanation of the survey—showing test forms and survey booklets—to the Sarpanch. Also, volunteers give the Sarpanch a form letter from the ASER Centre director that, beyond requesting support, invites the Sarpanch or "any other local citizen to join us in surveying the children" (ASER Centre, 2014c, p. 1). The meeting with the Sarpanch is an important step in the data collection process for ASER because, although the ASER volunteers do not technically need *permission* from the Sarpanch to conduct the survey, it would certainly be uncomfortable for them to proceed without support. According to ASER Centre (2014b),

> Village sarpanch [sic] is like the head of the family. Therefore, just like when one visits an Indian family we have to meet the eldest person first, it becomes important to seek the Sarpanch's permission to carry out the survey in the village…. It also helps us in gaining legitimacy in the village. (p. 7)

In addition, the Sarpanch typically provides information about the population, layout of the village, and number and location of government primary schools. Furthermore, the Sarpanch may enlist an assistant from the village or another member from the Panchayat to help the volunteers map the village or meet the government school headmaster. From my observations, some Sarpanches are talkative and enthusiastic about their village being selected for ASER while others are less interested in the survey, but I have never seen a Sarpanch refuse to let the volunteers proceed. In some instances, the Sarpanch is away from the village during the survey and so the volunteers ask around for a mobile number to gain support and information from the

Sarpanch via phone conversation or the volunteers meet with another member of the Panchayat who is available on that day.

Conducting the School Survey

The volunteers next walk to the government primary school, where they will conduct the school survey.[55] Occasionally, there is not a government school located within the borders of the village and thus, no school is surveyed. For reliability reasons, the volunteers must be careful to select the right government school according to ASER protocols when there are multiple ones within the village.[56] Once the volunteers arrive at the school, forging a positive and cooperative relationship with the school headmaster or head teacher is crucial. Volunteers must patiently and carefully "explain the reasons for collecting data on schools," which are "to see how much has been done on the front of parameters listed out in the Right of children to free and compulsory Education Act [sic], 2010," i.e., the RTE Act (ASER Centre, 2014b, pp. 15-16). If headmasters feel that ASER volunteers are conducting an inspection or that volunteers will ultimately "blame" the school staff for issues observed during the survey, it is likely the headmaster will not cooperate by not granting permission to conduct the survey, not providing access to school registers (i.e., student enrollment information), or not relaying other important information about school expenditures and governance. ASER Centre (2014b) cautions its staff and volunteers on this point:

[55] According to ASER training materials, the volunteers next complete village mapping, not the school survey. However, in reality, many volunteer pairs go to the government school after meeting the Sarpanch in order to ensure they can survey the school before it closes for the day, which often is earlier on Saturdays. This is an important adjustment to the survey process as the school will not be open the following survey day (Sunday). When the survey occurs over two weekdays, this is less of an issue. The other factor is how early or late the volunteers reach the village, which is highly dependent upon mode of transportation and distance/remoteness of the village.

[56] ASER Centre instructs volunteers to pick the government school with the most primary grades, i.e., an "upper primary school" with grades 1-8 versus a just a "primary school" with only lower grades of 1-5. However, the surveyed school must have at least lower grades 1-4 or 5 to be selected for survey. If multiple government schools in the village have all primary grades, the volunteers are to pick the government school with the highest enrollment.

> We should not make the Head Master feel threatened in any way. Also, while collecting the data, one should never start blaming him/her for the gaps and loopholes. Lastly, we have to stress that we are not a government body, and their or their school's name will not be disclosed anywhere in the report or to any government body. (p. 16)

The promise of confidentiality and the headmaster's belief that the volunteers are being truthful on this point is vital to whether the survey proceeds. Unlike with the Sarpanch, I have observed headmasters or teachers refuse to comply with the school survey, probably due to fear and a belief that their school is not operating well along certain parameters or according to RTE stipulations. With that said, I have also witnessed seasoned ASER state team members or MTs persuade an initially reluctant (or even somewhat hostile) headmaster into acquiescing to the survey and ultimately, to allowing the ASER volunteers full access for the necessary observations while also offering all the verbal information or documentation needed for each school survey indicator. When it comes to completing the school survey, the communication and persuasion skills of the surveyor can be the key difference between collecting or missing data.

Once allowed to proceed, volunteers collect survey data on the school from three sources: observation, documentation, and discussion (see school survey form in ASER Centre, 2015, pp. 42-43). Volunteers observe the number of children in attendance for each grade (via headcount) and the number of teachers teaching that day. The volunteers also observe whether the mid-day meal is being cooked (or was served) and examine facility-related indicators of school quality, such as whether there is drinking water available of drinkable quality and whether there are functional and reasonably clean girls' and boys' toilets. As far as documentation, volunteers ask to see the registers to note down official student enrollment figures for each grade level. Finally, volunteers sit with the headmaster or senior teacher to ask important questions about 1) knowledge of mandated student assessment practices under "Continuous and Comprehensive Evaluation" as required by India's Right to Education Act of 2009 (Government of India, 2009),

2) the operation of mandated school management committee (i.e., SMC), which includes village parents, and 3) receipt and use of government-allocated grants. The volunteers mark the indicators on the school survey according to verbal responses.

Mapping and Surveying the Village

The next step in the village data collection process is mapping and surveying the village. The volunteers will walk through the entire village observing its features while sketching a "rough map" that indicates its layout: the placement of homes, footpaths, and roads; the arrangement of its hamlets; and the location of all important landmarks like schools, banks, post offices, internet cafés, or health clinics (ASER Centre, 2014, pp. 2-6). For large and widespread villages, mapping can be a confusing and laborious endeavor; hence, ASER Centre encourages volunteers to "talk to people," "tell people about ASER," and "ask the children to take you around the village," (ASER Centre, n.d., p. 4).[57] Asking simple questions of villagers—such as "how many different hamlets/sections are there in the village?"—can save the volunteers considerable time while eliminating errors in the final map.

In my observations of ASER across multiple states, I saw volunteers typically receiving substantial help from the community in drawing maps. During the map-making exercise, it is commonplace for volunteers to have an adult or two accompanying them as well as a stream of inquisitive children gamboling behind. Because ASER volunteers and their activities are novel sights in these rural communities, they generate substantial curiosity amongst the old and young. The less timid villagers stop volunteers, ask questions, and offer their opinions on the map's depiction of the environs. More often than not, mapping the village is a spontaneous, amusing,

[57] The size of villages in India can vary immensely. For example, the 2011 Census of India for the state of Bihar recorded several villages with over 8,000 households and with less than 10 households. Because ASER Centre uses probability proportional to size (PPS) sampling method, larger villages are more likely to be sampled for each district. For more information, see ASER Centre (n.d.).

and communal process that sparks chatting across households about what these outsiders are doing and what ASER is.

ASER mapping is also an important technical exercise since volunteers convert this rough map into a final map that helps them divide the village into parts for randomly sampling households. Thus, if volunteers incorrectly draw the borders of a village—for instance, if they miss a remote hamlet of the village that is not easily found on foot (and consequently they do not realize it exists)—their omission (i.e., mistake) is compromising the representativeness and thus accuracy of ASER data.[58] Because the layout of villages can either be continuous (i.e., with consistently spread households in close proximity to one another) or in sections (i.e., with hamlets of households grouped apart from one another), ASER Centre has devised different protocols for delineating parts of the village on the final map from which to sample households (see ASER Centre, 2015, pp.46-48). In the easiest scenario, the village is continuous and the volunteers divide it into four roughly equal parts (geographically) and "assign each section a number," writing it on the final map (ASER Centre, 2014, p. 4).

During or by the end of the mapping exercise, volunteers must also fill out the Village Information Sheet. This form records the surveyors' names, date of survey, and village name along with information regarding the village's features, including a) whether it is accessible by a paved road, b) if it has an electricity connection or facility for solar energy, c) how many and what types of schools it has along with the grade levels they serve, and d) if there are post

[58] In some states (e.g., Bihar, Rajasthan and Uttar Pradesh), hamlets are separate living areas of the village that are divided by caste or religion and consequently, omitting hamlets is an important issue. While I was observing ASER field piloting in Bihar, our data collection team discovered that the lower caste hamlets were sometimes on the outskirts of a village and did not have the same amenities (i.e., electricity connection) as the other hamlets. Hence, inhabitants of some hamlets—perhaps, hamlets easiest to overlook during mapping—can have quite disparate living conditions from others.

offices, banks, health clinics and other amenities in the village. After the final map is drawn and the information sheet is completed, the volunteers are ready to sample households.

Sampling households. The volunteers must sample twenty households in total across the parts of the village that they have delineated (based on ASER protocols). Thus, in a village divided into four parts, the volunteers will need to select five households from each. The volunteers begin sampling households by walking to the center of one of the parts. In order to approximate a random sampling of households, the volunteers will select the first household on the left-hand side from the center of that part. Then, volunteers will proceed to sample every fifth subsequent house on the left-hand side (i.e., skipping four houses in between) until they have surveyed five households total in that part of the village. Next, volunteers walk to the center of another part and repeat the process until they have finished twenty households across all parts.

On paper, the sampling process appears simple. It is not always so in practice. For example, if a part of the village has households peppered on the side of a steep hill, it can be challenging to get a view of how they are arranged. Or, if households are not arranged in a linear way or if there are no clear paths connecting households, it can be difficult to establish what counts as a household "on the left." Or, when a volunteer is following a clear footpath and looking toward the left, occasionally there are households set behind other households. In such cases, the surveyors must make a decision: should those houses (set just behind other houses) get counted in selecting the fifth household? These extenuating circumstances in sampling are discussed during training sessions and ASER veterans will give newer staff members, MTs, or volunteers advice on how to deal with such decisions. However, such decisions are, in the end, judgment calls: they rely on the surveyor to interpret the purpose of ASER's sampling practice, which is to achieve a systematic sampling of households that is not tinged by convenience or

some other form of preference. Hence, the volunteers are charged with doing their best to approximate randomness in their selection of households.

Two other issues related to sampling include closed households or refusals to be surveyed. In the first case, no one may be home at the household selected for survey (or no one is willing to signal that they are home). Volunteers will approach the doorway of a selected home or the entrance of a household's yard (when the home has some kind of boundary fence) and call out. Sometimes, someone is visible in the doorway or elsewhere on the premises, but at other times, the home is clearly padlocked and shuttered. In such cases, the volunteers proceed to the next immediate household on the left. The other issue is when a household refuses to participate. Remarkably, this is less common than one might think. In general, as I described above, villagers are friendly to the volunteers, curious about their activities in the village, and generous with their time and information. Nonetheless, if households do refuse to participate, the volunteers sample the next immediate household on the left.

There are factors that can make refusal to be surveyed more likely. One factor that is within volunteers' control is their demeanor toward people in the household and their willingness to communicate the purpose of ASER to families. Typically, volunteers are young, and some come from city centers and have little experience in remote areas or with rural people in their districts. Their confidence and desire to communicate what ASER is matters, but their capacity to do so, particularly with older members of the household, may be limited. There are occasionally language barriers between volunteers and villagers: Villagers may speak a different dialect or a less common tribal language that the volunteers do not know. (In such cases, a younger and more educated person in the household may be able to translate or assist the volunteer with collecting the information needed from the household.) An issue beyond the

control of volunteers that can influence willingness to participate in the survey is gender. In particular, if both volunteers are men and the only people present in the household are women, social custom or fear can induce the women to refuse to respond to the volunteers' calls. Women may decline to admit male volunteers into their household or yard, or they may refrain from communicating with male volunteers at length. ASER state teams and partner organizations are frequently familiar with gender-related issues in the districts and may be able to anticipate these issues with surveying, but they may not be able to ameliorate all issues in the assignment of volunteers because of other factors, notwithstanding a paucity of female volunteers.

Completing the Household Survey

Once volunteers—armed with their ASER badges and survey packets—arrive at a sampled home, they ask to speak to its head or another adult living there. In some instances, an older teenager (i.e., 16-19 years old) may serve as a translator or as the actual household respondent (if parents or other family members are not present). After explaining in general terms what ASER is, the volunteers ask to proceed with the survey. The survey may be completed in a yard or on a footpath outside the home, at the home's entrance or on its front verandah, or inside the home itself. Members in the household determine the precise location of the survey. Commonly, household respondents invite volunteers to sit on a piece of furniture or they ask children to fetch plastic chairs or grass-woven stools from neighboring households (if the home does not own any). Villagers often make great gestures of hospitality toward volunteers, offering water, *chai*, biscuits, *namkin* (salty snacks), or meals as volunteers catch their breath, wipe sweat from their brows, and begin their routine of asking questions from the household survey form.

The ASER Household Survey Form collects essential information on three key areas: 1) the household's assets, 2) the parents and their educations, and 3) the children and their educations. Household assets are gauged as much as possible through observations made by volunteers, but some indicators require responses to questions. For instance, volunteers can observe if the home is *pucca* (i.e., made from concrete, bricks, timber and galvanized iron), *kutcha* (i.e., made from mud, bamboo, thatch, grass, and loosely packed stones), or *semi-kutcha* (i.e., made from a combination of pucca and kutcha materials). However, based on their location outside or within the home, they may need to ask if the family has electricity, toilets, a television, cable, or a motorbike/scooter. Volunteers also ask if the family has reading materials around the house (i.e., newspapers, magazines, and books). With regard to parents, volunteers record the names, ages and formal education levels of all parents (of children who are living in the household). If any other members of the household (not recorded as "parents") have completed 12^{th} grade, it is also recorded. If anyone in the household knows how to use a computer, it is marked on the form too. Finally, the volunteers enter into the form information about all children who are ages three to sixteen years and living in the household: the child's name, gender, age, enrollment status in school, language of instruction, and which type of school attended (i.e., government, private, madrasa, or other/alternative school). If children are too young to attend school, or if they are school age but have never been enrolled or dropped out of school, these details are also marked down by volunteers, including the child's grade-level at dropout and the year (e.g., 2005). The household survey also records whether children "take private tuition" (i.e., pay for private tutoring) and the amount paid monthly for tuitions. After recording the mobile number of someone in the home, the only portions of the Household Survey Form unfilled are

pertaining to scores for the ASER learning tests in reading and math. ASER 2014 also included an English language test that utilizes the same approach and format as the other two tests.[59]

Testing Children in Their Homes

To all children, ages 5-16 years old, who are living in a sampled household, ASER volunteers are supposed to administer the same basic three tests in the following order: reading, arithmetic, and English.[60] Volunteers have four versions of each type of test so children within the same household do not repeat each other's answers.

The objective of the ASER tests "is to find out the highest level that the child can do comfortably" for each subject area (ASER Centre, 2014a, p. 12). In support of this objective, the following are four vital issues covered in ASER's testing protocols:

1. **Testing environment.** The volunteer should ensure the environment is one in which the child can easily take a test and reasonably concentrate on the required tasks. For instance, the environment should be relatively quiet, and the child should be able to sit comfortably next to the volunteer with appropriate light in order to see the testing tool clearly. There are frequently barriers to an ideal testing environment that volunteers must actively mitigate: "Often family members and neighbors gather together to watch how the child is performing. This can make the child nervous." (ASER Centre, 2014a, p. 12). In particular, crowding can distract the child from the test's tasks and create undue anxiety.

[59] English language tests have also been given in some previous years of ASER (i.e., 2007, 2009, 2012). Given that "English-medium" private schools abound, that demand for English language skills in India is soaring, and that state governments are increasingly altering government school curriculum to accommodate more English language instruction, testing children on their English skills in addition to their reading skills in the regional language is often something that captures the interest of parents and state government officials alike. English is a high-priority subject.

[60] Even though the tests have a ceiling of a second- or third-grade proficiency, they are administered to children up to 16 years old. The logic is that many older children for various reasons—such as early school dropout, poor school attendance, or poor primary instruction—have not progressed beyond an early grade in their skills, so giving the basic tests to older children can also illuminate the degree to which they have struggled to "learn the basics."

2. **Interference.** Related to the issue of crowding is people's conduct while the volunteer administers the tests. The volunteer can not allow other children or adults to interfere in the testing process by touching the child, telling the child answers, or by scolding the child (for wrong answers or the inability to complete a task). A major obligation placed on volunteers is to shield the child from pressure so the child can demonstrate highest ability. ASER Centre recommends that good antidotes to interference are engagement and distraction: "One of the surveyors can talk to the adults or do some activities with the other children, while the other surveyor assesses the child" (ASER Centre, 2014a, p 12).
3. **Volunteer demeanor.** In addition to constructively managing family members and neighbors during testing, volunteers must form a good rapport with the child being tested. Thus, volunteers need to project friendliness and kindness. ASER Centre advises volunteers to have "a friendly conversation with the child before you start assessing them" (ASER Centre, 2014a, p. 12). Once assessing the child, the volunteer should "encourage the child by appreciating the effort she is making" and be patient if the child is struggling or slowly "reading or solving arithmetic problems" (p. 12). Volunteers' verbal affirmations of children's good efforts and smartness can solicit children to persist through tasks that initially intimidate them but that they ultimately can do.
4. **Second chances.** Volunteers must give children ample time to think and read. ASER Centre encourages volunteers to allow children to get familiar with the tasks on a test and allow children second chances to demonstrate their abilities: "To establish the highest level at which the child can comfortably do tasks, you may need to take the child through a series of tasks until you decide the level at which she really is" (ASER Centre, 2014a, p. 12). Many children will be unfamiliar with the oral administration and visual format of

the ASER test, so it can take them a small period to adjust, which merits allowing the children to return to harder tasks again later: "Practice and familiarity with a task improves the child's performance" (p. 12).

Reading test. The reading test is supposed to be administered in a child's "first language" (ASER Centre, 2014a, p. 12). For ASER 2014, ASER Centre produced the reading tool in over 16 Indian languages including Assamese, Bengali, English, Garo, Gujarati, Gurmukhi, Hindi, Kannada, Malayalam, Marathi, Meitei, Mizo, Nepali, Odiya, Tamil, and Telugu.[61] Nonetheless, it is not always possible to test children in their first language as some are from families who are linguistic minorities (i.e., who speak a non-dominant language in their district).[62] Due to India's immense linguistic diversity, ASER does not have reading tests translated into the mother tongues of many children.[63] When the ASER reading tool is not available in a child's first language, the child is usually tested in the state's prescribed language of instruction or in English (in the scenario that the child attends an English-medium school).

[61] This list is taken from the ASER Centre web page with links to the ASER 2014 reading testing tools in different languages (http://www.asercentre.org/p/141.html), but the list is missing the test in Urdu. Other languages may also be missing as one of the ASER Centre directors mentioned that ASER 2014 tested in "20, 21 languages" (Bhattacharjea, May 15, 2015).

[62] Language is a common issue in the administration of ASER in districts, which have a high number of tribal communities. Tribal people, called *adavasis,* are linguistic and ethnic minorities, who generally live in remote areas and sometimes observe nomadic livelihoods. Adavasi communities exist throughout many parts of India.

[63] In 2013, a study conducted by the Bhasha Research & Publication Centre revealed that approximately 780 languages are spoken in India, though many languages are threatened with extinction (see http://blogs.reuters.com/india/2013/09/07/india-speaks-780-languages-220-lost-in-last-50-years-survey/). This figure conveys why it is impossible for ASER reading tests to always be administered in children's mother tongues.

Figure 7. Surveyors conducting ASER reading test in a village household. "The Simplicity of Impact," by ASER Centre, 2014, retrieved from http://aserblog.com/wp-content/uploads/2014/08/Surabhi2Bpic.jpg. Copyright 2014 by ASER Centre. Reprinted] with permission.

Figure 8 depicts the basic decision tree that underlies the administration of the ASER reading test and determines the sequence of tasks a child is asked to complete by a volunteer. There are four sections to the reading test (pictured in Figure 1) that correspond to four reading tasks: 1) identifying letters, 2) reading single words, 3) reading four sentences (i.e., paragraph), and 4) reading a story (i.e., short passage). The volunteer always begins the reading test with the paragraph section by folding the testing page to only reveal the paragraph, ensuring the child is not distracted by the other test sections. The child is then handed the folded paper and asked to read the first sentence of the paragraph aloud to the volunteer. If the child can read all four sentences of the paragraph fluidly, with ease (even if slowly), and with three or fewer mistakes, the child has successfully completed the task for the paragraph section. The child is then asked to move onto the harder section with the story passage and to read it aloud (similar criteria determines successful completion of the story section). If the child reads the story successfully, the reading test is finished and the child is marked down as being at "story level"—the highest

proficiency measured by the ASER test. However, if the child cannot read the story after a second chance, she or he is marked as being at "paragraph level."

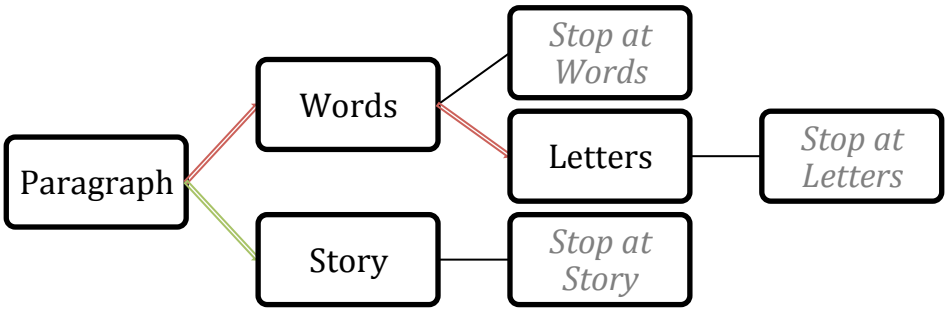

Figure 8. ASER reading test decision tree for tasks. The green arrows indicate the child got the previous task correct and is thus moving upward in difficulty in tasks. The red arrows indicate the child got the previous task wrong and is moving downward in difficulty in tasks.

Alternatively, if the volunteer administers the reading test to a child, who during the first task reads the sentences of the paragraph haltingly, only as a string of words, or with greater than three mistakes (after a second chance), the child is asked to read single words from the words section (as is the case if the child cannot read the paragraph at all). If the child can read (with ease) four out of five words, the test is finished. If the child cannot read words, the child is asked to identify single letters from the letters section of the test. Ultimately, the child is marked down as being at either "story," "paragraph," "words," "letters," or "beginner" level. The child is designated at the highest reading level completed with proficiency. For instance, if the child successfully reads the paragraph but not the story, the child is paragraph level. If the child cannot read most words and only identifies letters with ease, the child is at the letter level. A child who cannot read at least 5 letters is marked at "beginner" level. At the conclusion of the reading test, the volunteer begins testing the child in math and then English.

Math test. The math test follows a process that is similar to the reading test. Figure 9 depicts the decision-making process that underlies the math test and stipulates the sequence of tasks that the child is asked to do by the volunteer. There are four sections to the math test (pictured in Figure 2) that correspond to four arithmetic tasks: 1) identifying single-digit numbers, 2) identifying double-digit numbers, 3) solving double-digit subtraction problems (with borrowing), and 4) solving single-digit by triple-digit division problems. The math test begins with the volunteer asking the child to pick two subtraction problems from the subtraction section and copy them down on the back of the household survey form. The volunteer then asks the child to identify the double-digit numbers that comprise each problem. If the child can recognize the numbers, the volunteer asks the child to solve the two subtraction problems on the form, showing each step. In the scenario that the child solves the subtraction problems correctly, the volunteer instructs the child to copy down two division problems to solve on the back of the household survey form showing all steps in the process (see Figure 9). Once the child has completed the division problems, the volunteer either marks the child at one of two levels: at "subtraction" level if the answers given for the division problems are wrong (even after the child is given a second chance to review them), or at "division" level if both division problems are solved correctly.

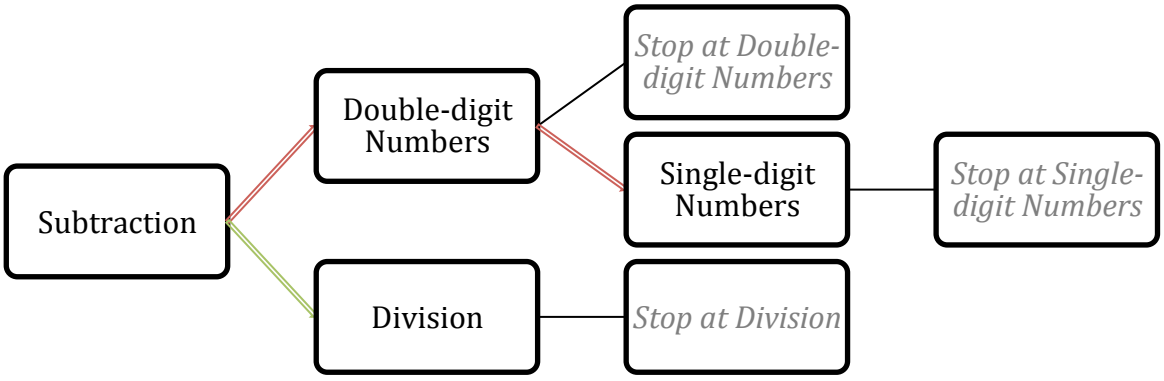

Figure 9. ASER math test decision tree for tasks. The green arrows indicate the child got the previous task correct and is thus moving upward in difficulty in tasks. The red arrows indicate the child got the previous task wrong and is moving downward in difficulty in tasks.

In the alternative scenario, the child cannot complete both subtraction problems correctly (after being given a second chance). Thus, the volunteer instructs the child to identify numbers from the double-digit number section. If the child cannot recognize the numbers, the volunteer asks the child to identify numbers from the single-digit number section. In either section, the child must be able to correctly recognize four out of five numbers to be proficient at that level. Depending on the child's proficiency, she or he is either marked down as at "double-digit numbers," "single-digit numbers" or "beginner" level.

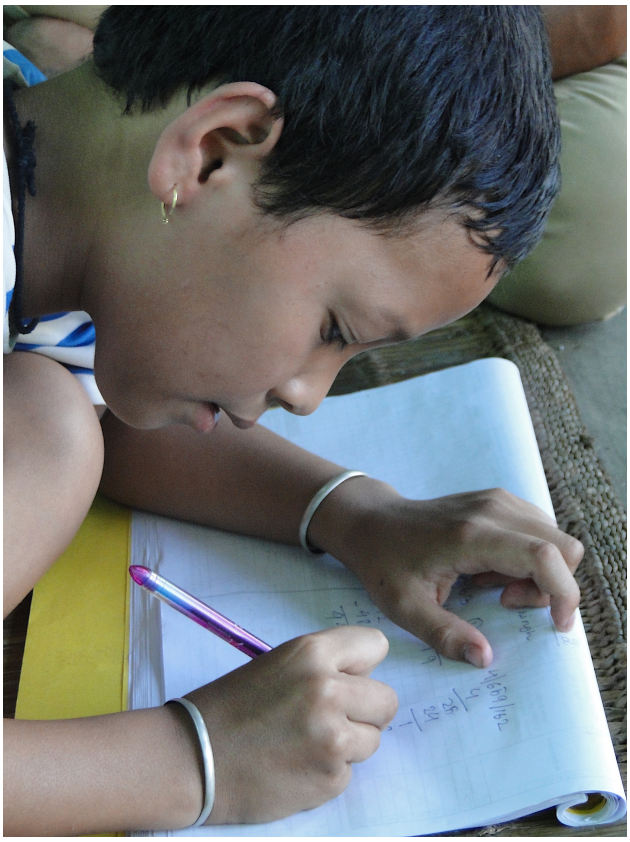

Figure 10. Child solving division problem for ASER math test. The child is writing down steps for completing a division problem as part of the ASER 2014 math test administered in households. Picture taken in a Manipur village, October 2014. Copyright 2014 by ASER Centre. Reprinted with permission.

English test. The English test has a format similar to that of the math and reading tests but its testing process is slightly different. Figure 11 is a picture of the one-page English test used during ASER 2014. It includes four sections that correspond to four English language reading tasks: 1) identifying capital letters, 2) identify smaller case letters, 3) reading single words and explaining their meaning, and 4) reading four sentences and explaining their meaning. Thus, children can be identified as having one of five different levels of English proficiency—"beginner," "capital letters," "small letters," "words," and "sentences."

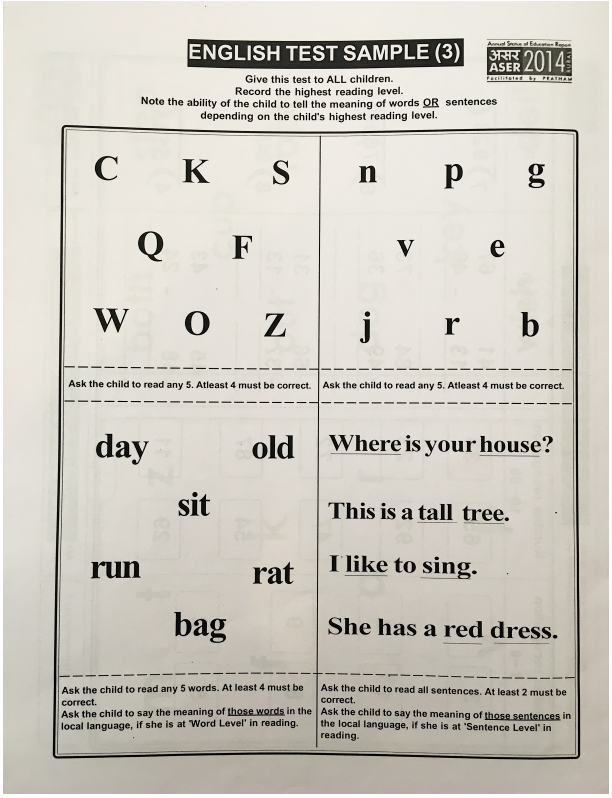

Figure 11. ASER 2014 English test (sample 3). One of the four sample English tests used in ASER 2014 taken from the official testing booklet used by surveyors in households From "Testing Tool," by ASER Centre, 2014, p. 12 . Copyright 2014 by ASER Centre Reprinted with permission

The decision-making process that underlies the English test and stipulates the sequence of tasks that the child is asked to do by the volunteer is depicted in Figure 12. Unlike the reading and math tests that begin with a task of mid-level difficulty, the English test begins with the simplest test task—recognizing capital letters—and moves progressively through more difficult tasks—first, smaller case letters, and then, words and sentences. The English test stops when the child cannot complete the requested task successfully. The other feature that distinguishes the English test from the reading test is that children are asked to state the *meaning* of words after being marked at "word" level and of sentences after being marked at "sentence" level—volunteers mark down the child as "can say" or "cannot say" meaning for words and sentences. Thus, the English test also assesses children's comprehension of language, not just their reading

ability. ASER Centre's logic for testing comprehension on the English test is that nearly all children who study English in school are learning it as a second, third, or fourth language—that is, it is not their mother tongue—and "children start learning English with capital letters followed by small letters" (ASER Centre, 2014d, p. 13).[64]

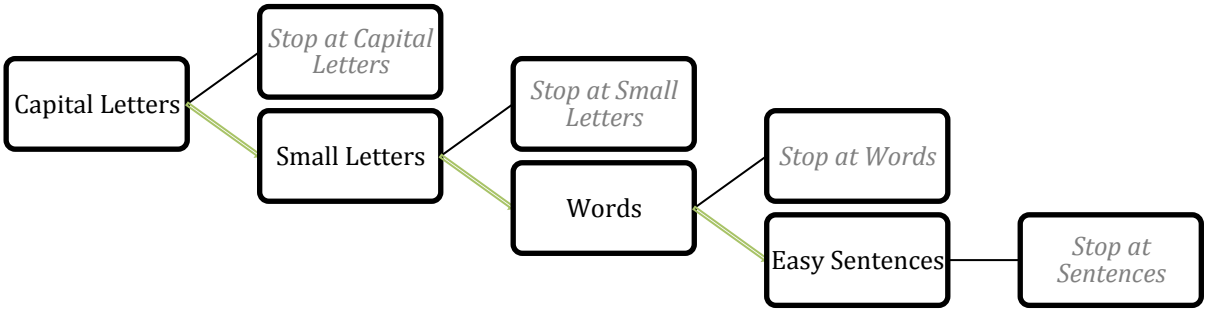

Figure 12. ASER English test decision tree for tasks. The green arrows indicate the child got the previous task correct and is thus moving upward in difficulty in tasks. (Otherwise, the test concludes with the previous task.)

Refusal to test. A major challenge for volunteers is when children refuse to take the ASER tests. It is not uncommon for children, especially younger ones, to feel apprehensive about participating in testing and to run away or hide from volunteers, family members, and other bystanders like school friends and neighbors. Even adolescents and teenagers occasionally lock themselves in rooms or flee on-foot when they discover that part of the household survey includes testing—their ingrained trepidation over exams quickly overcomes any curiosity about

[64] With regard to testing reading comprehension for children who are being tested in a language that is not their mother tongue, there is a bit of contradiction between the ASER reading and English tests. In reality, during the ASER reading test, children of linguistic minorities are being tested in a regional language that is not their first language, but the reading test does not assess their comprehension of what they are reading. ASER Centre has relayed that in years past where they have taken the additional step of testing comprehension in addition to reading on the reading test, they have not obtained significantly different results in terms of children's observed learning levels and thus, testing for comprehension does not seem to be a useful additional step in ensuring accurate data.

these strange volunteers, which may have kept them lingering around. Sometimes, parents, siblings or friends are persuasive in convincing the child to come out, come back, calm down, and test, but at other times, their tactics of persuasion are anything but persuasive (or constructive) for the child. Parents and grandparents can especially take a child's refusal to test badly—harshly scolding or manhandling a child in order to try to force him or her into testing—hitting, holding down for testing, or dragging the child by the most catchable limb to where the volunteer is waiting. To be clear, ASER Centre and ASER volunteers do not encourage adults to force children into testing. In fact, the Centre's leadership recognizes that such tactics do not align well with ASER's espoused values and they do not enable children to do their best during testing. When children are intimidated into participating in testing, it can adversely affect the reliability of ASER learning data.

Nonetheless, too many refusals from children to do testing also presents a long-term issue for the quality of ASER's learning data—in short, there is no guarantee that those children's missing scores on the reading, math and English tests are equivalent to the children who are willingly sitting for testing. Thus, volunteers are encouraged to do what they can to convince children to complete testing and to put the children at ease, so volunteers can observe them completing tasks at their highest ability level. Persuading a reluctant child to sit for testing is one scenario during the ASER data collection in villages that distinguishes confident or experienced surveyors with good communication skills from those who are less adept at the ASER process, and particularly, who are less comfortable interacting with children and their rural families. It is in such circumstances that the wisdom of ASER Centre's approach becomes apparent: specifically, their guidelines for pairing strong volunteers with weaker ones, and for encouraging master trainers to monitor their weakest pairs of volunteers. Such measures can ameliorate this

testing challenge. In the latter situation, master trainers can step in if volunteers are having difficulties and try his or her luck at establishing rapport with the child and administering the tests.

Monitoring of Data Collection Process

Since 2011, monitoring and recheck activities have been "an integral part of the ASER Process" as the key components of its "quality control framework" (ASER Centre, 2014e, p. 1).[65] There are two stages to monitoring the work of volunteers: pre-survey and field.

Pre-survey. The first stage, briefly mentioned under the district-level training section, is the monitoring that takes place "pre-survey," that is while volunteers are being trained. This first stage of monitoring entails the master trainers evaluating the volunteers' performance in three areas: 1) on the quiz that assesses volunteers' understanding of the whole ASER data collection process (including the proper protocols for testing); 2) during observations of volunteers throughout the field-based piloting of data collection in villages; and 3) based on the volunteers' attendance throughout the training sessions.

Field. The second stage of monitoring is during actual data collection in the field; it can be in-person or by phone. With regard to phone monitoring, MTs are directed to call all their pairs of volunteers on the days of data collection to inquire about how the process is going. MTs should offer encouragement or further guidance if volunteers encounter dilemmas or have questions while in villages. As for in-person monitoring, most pairs of volunteers are alone

[65] ASER Centre also introduced another layer of monitoring in 2011 called the ASER Call Centre "where each state hired a call centre leader to routinely make calls at each stage of the survey period to track survey progress and flag any problems" (ASER Centre, 2014d, p. 5). For a more detailed discussion of the evolution of ASER monitoring and recheck practices as well as more information about the Call Centre, see ASER Centre (2014d).

during their data collection in a village, but some pairs have a master trainer or ASER state team member monitoring them.

As discussed in the last section, the MTs are instructed to in-person monitor certain volunteers, depending generally on two criteria: 1) the pair of volunteers is "under confident" or relatively weaker than the others (ASER Centre, 2014d, p. 4) and 2) the volunteers have been assigned a village that is comparatively "difficult to reach" because of its remoteness, difficult terrain, or sheer distance from the partner organization or town center. Occasionally, a village is difficult to reach because of extreme weather conditions or due to security concerns in the area. In the last instance, in addition to monitoring them and depending on the severity of the security concerns, ASER state team members or seasoned MTs may assign a pair of volunteers who they think for linguistic, cultural, experiential, or other reasons are going to encounter the least amount of problems during data collection. However, there are instances where security concerns prompt the MTs or ASER state team members to survey the village themselves (i.e., without volunteers). As a last resort, ASER state team members may request that ASER Centre headquarters eliminate a village from the sample and replace it with a new village.

The proportion of volunteers from a district who are monitored in person depends on how many volunteers were recruited and trained for the district and how concentrated the period of data collection is—a lengthier data collection period with less volunteers enables MTs to monitor a higher proportion of the district's volunteers (see Appendix F). Generally, data collection in a district happens over a period of one or two weekends.[66] Regardless of the number of volunteers

[66] If 60 volunteers are successfully recruited and trained in a district, there are enough volunteers to assign a pair to all 30 villages and thus, data collection can be completed in one weekend. However, if there are fewer volunteers for a district, a pair of volunteers may need to collect data for multiple villages over successive weekends. The disadvantage of fewer volunteers is it takes longer to complete data collection for all the district's sampled villages; however, the advantage is a greater proportion of volunteers are monitored in person at some point, thereby logically increasing the reliability of the data for the district.

in a district, a two-weekend data collection period allows for double the number of villages with oversight (monitoring) of their data collection. If the ASER data collection occurs over two weekends, one MT can monitor one pair of volunteers in a village per day and therefore, two pairs in two different villages during one weekend. Accordingly, over four days of surveying (two weekends), between the two MTs assigned to the district, they can monitor a total of eight volunteer pairs (16 volunteers total) or eight of the 30 sampled villages for the district. ASER state team members also act as additional monitors, increasing the proportion of volunteers monitored in some districts. MTs and ASER state team members carry with them formats to fill out, which help guide what needs to be monitored. Monitors check "yes" or "no" boxes in response to the form's list of questions including: "Has the map been made properly?" "Are the surveyors talking politely to the respondent?" Did the volunteers carry with them the "Tools (Language, Math & English)"? (ASER Centre, 2014d, p. 6). While monitors are filling out the forms, they are supposed to also "ensure the survey [i.e., data collection] is being carried out as planned and any mistakes being made by surveyors [i.e., volunteers] are being rectified on the spot" (ASER Centre, 2014d, p. 4). Thus, monitors have both a monitoring and an intervention function—they are charged with correcting data collection mistakes and re-teaching the appropriate protocols to volunteers, ensuring the remainder of the data collection process is executed properly.

Rechecking and Compiling Data, and Resurveying (Phase IV)

As mentioned, the other part of ASER's quality control framework is recheck. Most recheck activities are quickly executed by MTs once volunteers have completed a round of data collection in villages. Three types of recheck activities comprise MT's responsibilities: desk recheck, phone recheck, and field recheck. In addition, MTs must compile ASER data and

analyze it for red flags during recheck. Depending on what is discovered during the three types of recheck and while compiling data, some villages, households, and schools may need to be resurveyed and children retested. The following sub-sections describe in detail the routine rechecks conducted by ASER master trainers. Furthermore, a second layer of recheck for ASER is conducted by ASER state team members and a final layer of recheck is undertaken for some districts (referred to below as "external recheck" and "ASER Centre Recheck").

Desk Recheck

As soon as volunteers are finished completing data collection in a village, MTs are encouraged to obtain from them the packets of survey forms for desk recheck. According to the "Quality Control Framework" training guide, "Desk recheck is the thorough checking of the survey booklets, received from the surveyors [i.e., volunteers], in order to identify villages where the data is incomplete" (ASER Centre, 2014e, p. 9). MTs should check the forms "in the presence" of the volunteers so that they "can be questioned if any missing information is found" (ASER Centre, 2014e, p. 9). In-person review of the data collection forms is especially beneficial in "the case of a two- weekend survey, where the surveyors will be surveying another village, [as] this helps to strengthen the surveyor's capability" (ASER Centre, 2014e, p. 9). MTs utilize a checklist created by the Centre for guiding their verification process of the data forms.

Phone Recheck

Following desk recheck, another aspect of recheck that verifies if the survey has been conducted properly is phone recheck: "the process by which MTs make phone calls to certain households in every village" (ASER Centre, 2014e, p. 11). MTs are responsible for randomly calling at least eight of the 20 households from each village to confirm the household survey was administered and children were tested. If more than one household denies that data collection

happened, the MT must call an additional four households from the 20 sampled (12 households total). In addition to calling randomly sampled households, MTs should purposively sample and call any households that during the desk recheck had information that appeared dubious in any way. The purpose of the phone call is to ascertain at a very basic level of quality if the survey and testing were conducted properly. As such, MTs should try to speak directly with the recorded respondent on the household survey form and ask direct questions that request information that will be easy for the respondent to recall, such as:

1. "Did 2 people come to survey your HH [i.e., household] on <mention the date>?"
2. "Did they ask information about the children (between 3-16 years) in the household?"
3. "Did they show any Language, Math and English questions to the children?" (ASER Centre, 2014e, p. 11).

Referencing the forms, MTs can use the recorded names and ages of children when asking the respondent the above questions. MTs are trained by ASER state teams to be "very polite" while speaking to household members, to "explain the purpose of the phone call very clearly," and to not be "impatient to receive the information you need" (ASER Centre, 2014e, p. 11). Furthermore, the Centre wants MTs to be sensitive to rural families' routines by not making calls "at a time that is… interfering with anybody's daily activities" or late at night (ASER Centre, 2014e, p. 11).

Compiling Data and Verifying Quality

Once they have done the desk and phone rechecks, MTs must compile data by filling out Village Compilation Sheets (VCSs) and the District Compilation Sheet (DCS). Together, these sheets aggregate all the data collected by volunteer pairs at the village level and then, the district level. For the VCS, MTs are required to "manually do the compilation of household data

themselves"—the Centre warns this must be executed "very carefully and correctly, because all other compilations are based on the totaled information in this sheet" (ASER Centre, 2014e, p. 12). Following the completion of a VCS for each of the 30 villages in a district, the MTs enter VCS information into the DCS, which "is a format that captures all data for the entire district" (ASER Centre, 2014e, p. 12). At this point, the DCS presents another opportunity for MTs to verify the quality of their district's data: the DCS "can…be used to analyze trends and study anomalies in specific villages" (ASER Centre, 2014e, p. 12). ASER state team members train MTs to spot strange patterns and trends in their village-level data that would indicate that volunteers either fabricated data or executed the survey and test process incorrectly or incompletely. For instance, if volunteers have recorded test scores for children that appear symmetrical for reading, math, and English, it is a sign that the testing data for that village may not be trustworthy. Another example is if that village's sampled households have on average many less school-age children than other villages in the district. This may indicate that volunteers did not survey households at the appropriate time to test children and subsequently decided to omit the children's information from the household survey forms (rather than have to return to the households to conduct testing when the children are home). In cases where data seem irregular, the Centre advises MTs that "it would be wise to conduct a thorough phone recheck in these villages to be sure of the survey quality. Field recheck can follow, if required" (ASER Centre, 2014e, p. 13).

Field Recheck

At this stage in its quality control process for ASER, the Centre states that MTs have three sources of information from which to make sophisticated judgments about "finalizing [which] villages [to choose] for field recheck" of data (ASER Centre, 2014e, p. 13), specifically:

1. Feedback from volunteers and data completeness from desk recheck
2. Feedback from households and reported problems from phone recheck
3. Worrisome trends and anomalies in data from analyzing the DCS.

If the desk and phone recheck processes did not uncover any concerns with data quality, the MTs are instructed to rely on their prior evaluations of volunteers during pre-survey monitoring to make suggestions for which villages are rechecked. Eventually, the ASER state team members (in consultation with MTs) determine which villages are field rechecked. According to ASER protocols, the two MTs for a district "together must [field] recheck at least four villages after a one weekend survey, and at least eight villages after a two-weekend survey" (ASER Centre, 2014e, p. 13). (See the last table in Appendix F for number and percentages that correspond to the general guidelines for the number of villages and households verified during each type of recheck.)

Once villages for recheck are chosen, MTs must carefully plan their travel to these locations and arrive early. MTs are required to complete a range of recheck activities in a concentrated amount of time:

> The MT must visit both, households and the school. For this, the MT must reach the village much before schools have opened or just an hour before schools close, so as to spend time with children in the households and also the HM [i.e., headmaster]/teacher in the surveyed school. (ASER Centre, 2014e, p. 14)

MTs carry with them the volunteers' survey booklet to be rechecked, the packet of ASER testing tools, the ASER 2014 recheck format, and an extra survey booklet—the latter "in case the MT finds that the village to be rechecked hasn't been surveyed at all or has been surveyed very badly she/he can begin resurveying the village her/himself" (ASER Centre, 2014e, p. 14). The master trainers are given a set process for field recheck. First, they must randomly select ten households from the survey booklet "where children in the age group of 7-16 years have been tested" but

ensuring that all surveyed sections of the village are represented (ASER Centre, 2014e, p. 14). Next, the MTs refer to the recheck format that requires old and new information from four sources: 1) the original household survey sheet filled in by volunteers, 2) the adult respondent's new responses to questions, 3) the tested child's new responses to questions, and 4) the MT's observations of the household during recheck. Importantly, MTs do not retest children who were tested by the volunteers. They simply ask the child questions about whether they were tested and in what sequence were they asked to complete tasks. The MT's job is to ascertain if the volunteers indeed tested the child and if the volunteers appeared to follow the correct testing process.

Field Rechecks and Monitoring by ASER State Teams

During the monitoring and recheck process, ASER state team members are mirroring many of the activities of MTs to add a second layer of trustworthiness to the ASER process and its data quality. State team members fulfill monitoring and recheck duties at a higher level of competency and importance than MTs as they are simultaneously evaluating the volunteers' and MTs' work. For instance, ASER state team members may be extra monitors of volunteers in a district that is difficult to survey, but they are also taking steps to ensure the quality of monitoring being done by MTs in that district. With regard to field rechecks, state team members recheck "some selected villages" based on "the performance of Master Trainers and the surveyors" (ASER Centre, 2015, p. 65).

External Recheck and ASER Centre Recheck

A final layer of field rechecks for ASER—"external recheck" or "ASER Centre recheck" (ACR)—are conducted for some districts (see ASER Centre, 2015, p. 65; ASER Centre, 2014a). For ASER 2014, external field rechecks were conducted by "colleges and NGOs across India"

for randomly selected districts and villages" in the states of "Assam, Gujarat, Haryana, Jharkhand, Karnataka, Madhya Pradesh, Mizoram, Tamil Nadu and Uttar Pradesh" (ASER Centre, 2015, p. 65). In terms of ACR, like external rechecks, it is executed without MTs. ACR is implemented via a large cross-state exchange of ASER state team members (with roughly 10-12 states participating every year). ASER Centre's central office staff tightly control the assignment of state teams for ACR and they determine what districts will be rechecked—some "chose[n] purposively" and others that are "selected randomly" (ASER Centre, 2015, p. 65). Then, state team members switch states and conduct extra field rechecks of ASER data according to the centrally defined plans. ACR is intended to contribute an additional layer of transparency to recheck processes and further "ensure that the same standard of survey [is] conducted across the country"—that is, the same standard of quality for every state (ASER Centre, 2014a, p. 25). Since ACR is the last layer of the quality control process for ASER, results from ACR are taken very seriously as it is indicative of the data quality for the entire district and state after all internal checks have been completed. During ASER 2014, "approximately 56% of villages surveyed…were either field monitored or field rechecked by Master Trainers, ASER State Teams and others" (ASER Centre, 2015, p. 65), and for some states like Uttar Pradesh, the most populous state in India, the number of surveyed villages either monitored or rechecked was over 70% (ASER Centre, 2014a, p. 29).

Resurvey

When ASER partnerships are formed, the organizations are made aware that they are responsible for ensuring that ASER data quality is good—that is, the district's data turns out to be free of major issues like missing, fabricated, or otherwise problematic data due to surveyors' or MTs' faulty adherence to ASER protocols. Otherwise, if during recheck the "survey or data

from a particular village(s) is found to be of poor quality, then that village will have to be resurveyed by the partner organisation at no additional cost to ASER Centre" per the partnership agreement (ASER Centre, 2014a, p. 33). In practice, however, ASER Centre may opt to not involve the partner organization in resurveying efforts. In the event that an ASER state team member (having worked closely with the partner) does not trust that partner's capacity or commitment to safeguarding the quality of the resurvey process, the ASER state team will likely decide to resurvey the village themselves or to send a trusted MT to do it rather than risk being fed bad data by the partner for a second time and thereby, delaying the completion of the district's data collection further. In short, if there is concern that the resurvey will be conducted improperly by the partner organization, ASER Centre headquarters would rather its state team pursue collecting trustworthy data more than recouping its investment of money and time in the partner.

At any time during the recheck process, ASER state team members in consultation with their respective state's Research Associate may conclude that resurveying a village is necessary. The most common causes for resurvey are volunteers dishonestly filling in data (frequently not even visiting the households or villages), or volunteers incorrectly administering the ASER testing process or wrongly sampling households. These are considered non-negotiable violations of the quality of ASER data. Nevertheless, ASER Centre directs MTs not to make a decision to resurvey a village or part of a village, school, or household on their own even if they encounter these issues: "If MTs face any problems during the recheck or find that the survey was not done properly, they must call up ASER Centre immediately" (ASER Centre, 2014e, p. 15). A decision to omit prior data and commit time and resources to resurveying the village is considered a crucial decision that should ultimately be made by the state and central teams, though MTs are

often an essential source of information regarding such decisions. Without MTs, ASER state teams frequently would not know that anything had gone awry with testing and surveying in a village. In sum, master trainers' ability to monitor volunteers and their competency with rechecking data remain ASER Centre's greatest line of defense against poor data quality.

Aggregating and Analyzing Data and Report Writing (Phase V)

When the data collection, monitoring and recheck processes are all completed for a district, the MTs send off the hard copies of the 1) survey booklets for the 30 villages, 2) the District Compilation Sheet, and 3) the attendance sheet for the district training. The data from these hard copies is "entered into the software of the data entry centre" for that state (ASER Centre, 2014e, p. 19). MTs are also instructed to send a scanned version of the DCS by email or a photocopy by courier to the ASER Centre headquarters in Delhi where central staff track district completion and double-check data. Aggregating data happens in ASER Centre-contracted data entry centers whereas analysis of data and report writing is undertaken by a small team at the ASER Centre headquarters in Delhi.

In terms of preparing the report, data analyses are led by Dr. Wilima Wadhwa and report writing is collaborative and shared amongst the three ASER Centre directors, Drs. Wadhwa, Rukmini Banerji, and Suman Bhattacharjea. An opening statement or article in the report is written by Pratham founder, Dr. Madhav Chavan, who was also a main author of the report during the first years of ASER. Introductory articles are penned by ASER Centre and Pratham directors and senior staff members as well as by important collaborators of the two NGOs and prominent experts who use ASER findings to investigate particular topics or discuss issues related to education and the history or influence of ASER (see ASER Centre, 2015, pp. 1-28). The beginning of the ASER report also includes a listing of all partner organizations who make

ASER possible in each state and a multicolored map that displays the type of partner organization (e.g., university, DIET, and NGO) for each district in each state. A large portion of the 2014 report's 335 pages is dedicated to offering clear explanations (with accompanying charts, diagrams, and artifacts) of how the ASER data collection process works with examples of the protocols and formats used to collect the data. An especially appealing aspect of the ASER reports are the captivating photographs depicting villages, households, schools, and children taking tests from all over rural India. Snapshots of ASER trainees and volunteers interacting with the community are taken from interesting angles and capture the action of ASER.

Over time, the final report has come to display standardized tables and figures from year to year that feature data for all of India (in the first section) and by each Indian state (in subsequent sections). Tables report data on children's learning levels in reading (regional language), math, and English language by enrollment type (e.g., government school, private school, and out-of-school), grade, age, and gender. There are also tables that address findings related to government schools' status on indicators for India's Right to Education Act, ranging from pupil-to-teacher ratio and number of multi-grade classrooms to provision of drinking water and girls' toilets. In the initial section for all of India, some tables compare "the performance of states" on learning levels disaggregated by private school enrollment or "the performance of schools" by state along RTE indicators (see ASER Centre, 2015, pp. 89, 94-95). Figures that are typically included in reports are multicolored maps that compare states' private school enrollment rates or learning levels—for example, the state's categorization by percentage of grade-five students who can read a grade-two story (see ASER Centre, 2015, p. 80). Thus, multicolored maps of India visually summarize ASER findings for reading levels, math levels, private school participation, and so on. Following the sections that report state-level data are

sections with divisional estimates for some states (depending on sample-size thresholds). Reports conclude with appendices that provide an explanation of ASER's sampling design and describe that particular year's sample.

Publishing ASER and National Media Release (Phase VI)

Hundreds of final reports with the ASER findings for that year are published by early January and publicly released by mid-month. There is a flurry of cover-letter printing and large-envelope stuffing in ASER Centre Delhi headquarters as central staff members' days are overtaken with sending out countless copies of the report and glossy executive summaries and special inserts. Reports are often accompanied by personally addressed letters mailed to all members of parliament, senior officials in government ministries, funders, partner organizations, and important ASER supporters domestic and international. For days, mailing materials occupy nearly every open surface of ASER Centre's main staff room, which includes the tables and desks that are workspaces for roughly 30 people.

The report is officially released at an annual ASER National Media Launch where ASER Centre directors and Pratham leadership and board members speak in addition to prominent dignitaries and supporters. ASER 2014 marked the ten-year anniversary of ASER, and thus the Centre hosted an especially large media release with panel speakers including Professor of Economics at Massachusetts Institute of Technology and founder of the Abdul Latif Jameel Poverty Action Lab (J-PAL), Abhijit Banerjee; Director of UNESCO's Global Education Monitoring Report, Dr. Aaron Benavot; and Professor of International Development, Lant Prichett of Harvard University's Kennedy School of Government. The 2014 national media release was also notable because it was attended by funders like the Hewlett Foundation, NGO collaborators in other countries like UWEZO (in Kenya, Uganda, and Tanzania) and ASER

Pakistan, and by some of the ASER 2014 partner organizations from across the country. Special funding was procured to bring a few partners from each state for the 2014 release and engage them in listening and discussion groups following the launch festivities. These groups were intended to generate new ideas about what ASER Centre and the ASER evaluation should be tackling in the future.

ASER Centre has relied heavily on media coverage—both newspaper and television—to disseminate ASER findings each year. The media coverage raises public awareness about learning and creates pressure on the national government and state governments to address issues with learning and school provision that are highlighted in ASER's findings, especially as ASER has grown in prominence. More recently, ASER is covered widely not only in Hindi- and English-language national media but also in state-level or local media reporting the findings in regional languages. *The New York Times* profiled ASER ahead of its 2014 release of findings and then included ASER 2014 data in an editorial about Indian schools a couple months later (Kingdon, 2015; Rosenberg, 2014). However, media coverage in India, especially relying on the media as the principal form of data dissemination to the public, has its issues for ASER Centre. Even with online access and print access to the actual ASER report, most people from Government of India officials to academics to ordinary citizens learn about ASER findings from national and local media coverage. Sometimes, the media wrongly represents how ASER data is collected, what ASER findings show, or how ASER findings should be interpreted and the implications of this data for school reform. As a result of these issues with media coverage, the directors have attempted to do more writing themselves for popular and academic consumption (i.e., newspaper articles and academic journal articles) and make media appearances (i.e., do television interviews), but the Centre's leadership expresses that it is difficult for them to

monitor and address all of the problematic representations and interpretations of ASER data. (ASER Centre currently does not have the capacity to do so with its small media and leadership team.)

Relatedly, the directors have expressed that there are ongoing challenges related to presenting quantitative data in a country where there is not a "measurement culture"—meaning there is not widespread understanding of how to "use large-scale evidence to inform decision-making at *whatever level*" (W. Wadhwa, personal interview, May 15, 2015). Consequently, not only are the media outlets poorly equipped to analyze ASER's assessment and quantitative data in light of how it can be used to inform understanding about the Indian education system, but according to ASER Centre, even among many social scientists (i.e., academics in Indian universities) this competency is quite low. Such a need for more developed understanding of how to use evidence in decision-making spurs ASER Centre's desire to build this culture of measurement within India through sharing ASER findings and via public deliberation of ASER findings' meaning. However, this need reflects a current atmosphere in India that presents unique challenges in terms of addressing claims about issues with ASER's credibility or quality. When individuals claim that ASER data are poor quality or that the ASER sampling frame is too narrow or problematically designed, the directors express that these sometimes politically motivated critiques cannot be address through technical arguments because they come, in part, from a lack of understanding about sampling and other issues related to quantitative data.

Wider Dissemination of ASER Findings (Phase VII)

Following the distribution of printed reports (via mail) and the national media release, ASER state teams plan state-level media releases or other events to share ASER findings with key officials and stakeholders. Dissemination activities at the state and local levels have been

much more variable than those at the national level. Overall, dissemination activities are less uniform than the clear protocols that govern the ASER process up until the report's production. There are several reasons for this variability:

1. **ASER's theory of change**: According to the "theory of change" that animates ASER's design (W. Wadhwa, personal interview, May 15, 2015), its findings (i.e., results) are meant to prompt change at the national and state levels through direct dissemination and use by policymakers and via public pressure as an outcome of media coverage (Goodnight, Paper 3). Consequently, the ongoing focus has been to disseminate the report to powerful people within the national and state policy-shaping communities and then generate a lot of media coverage. In contrast, ASER's influence at the district level and with partner institutions, communities, and volunteers who participate in ASER is supposed to come via the ASER process and the awareness and capacities that it builds.
2. **Resource constraints**: Resource constraints are another reason that dissemination plans are not more robust for each state. It requires substantial manpower, resources, and time for ASER state team members to return to partner organizations and villages to disseminate findings and engage in discussions about possible solutions. Historically, ASER Centre has not had the funding to execute this kind of engagement around the findings. Therefore, partners, communities, and schools have been providers of ASER findings, but by and large, have not been consumers of those findings.
3. **Experimentation**: On occasion, ASER state teams have experimented with different approaches to dissemination by holding an event or facilitating a workshop to share ASER findings with a particular partner organization or community (usually due to expressed interest in the results and in planning an intervention). For example, for ASER

2014, the Centre received special funding to do this type of dissemination and further capacity-building work with a handful of partners. Unfortunately, the results of these dissemination activities are not part of my study's data.

4. **Pratham:** Another reason that the Centre has not engaged in more grassroots dissemination of findings is that there was an ongoing assumption that Pratham would create the solutions and interventions based on ASER findings in collaboration with communities and schools. However, ASER Centre directors worried that there was a conflict of interest in ASER Centre being involved in both data collection and the advocacy of particular types of solutions and interventions. This concern has more recently faded as ASER Centre has embraced the idea of being involved in providing solutions in response to constant requests from government officials and partners for help: ASER data continues to show flat if not declining learning levels in children across rural India.
5. **Individual and collective initiatives**: Finally, ASER was designed according to its theory of change—to inspire self-directed individual and collective grassroots initiatives as a response to participation in the ASER process. There is a hope underlying ASER that its volunteers and partners will be inspired to work for the cause of education long-term and take personal responsibility for effecting that change wherever they see the opportunity to do so in their immediate spheres of influence.

Individual and Collective Grassroots Efforts

ASER Centre hopes the ASER process has a lasting influence on its participants' capacity and commitment to be engaged in education efforts moving forward. In the immediate

future, the Centre encourages partners and volunteers to discuss the ASER findings with people in their constituencies, communities, and spheres of influence to generate widespread awareness and debate about primary education and learning. With regard to long-term objectives, the real aim of ASER is to motivate use of its findings and action beyond just the months immediately following the report's release.

One challenge for the Centre is in trying to stay abreast of the impact of the ASER process and its findings as individuals return to their communities and lives after the annual evaluation concludes. Fazal, an ASER manager from the state of Jammu and Kashmir, wrote the following blog post entitled "Life Lessons" that illustrates the confidence, passion, and skills that ASER instilled in him and the lasting influence that it had on his educational efforts: [67]

> Prior to joining ASER Centre, I was shy and had great trouble speaking in public, be it with a group of friends or a small gathering. I was so shy that I could not talk to my relatives and neighbours without getting self-conscious. I have been working as a teacher in a Madrasa (Islamic school) in Kashmir for a few years now but there was something lacking in the way I was going about my job. And that was 'exposure'. Four years ago, I joined ASER Centre as an ASER Manager to support the team in conducting the annual survey in Kashmir. I had never left my hometown until this point. But all that changed, and I got to travel to other parts of Kashmir and beyond to conduct the survey. But the real learning came in the form of life lessons. ASER has taught me discipline, punctuality, teamwork, planning, communication, etc., and I try to apply these qualities in my full-time job too. Two year ago, following a discussion over the Annual Status of Education Report findings with fellow teachers, parents, friends and local community members, we formed a group called Karwan-e-Tableeg Al-Quran (KTQ) Meer Behri Dal Lake. The objective of this initiative was to support students and schools in the district with class work, conduct exams, and provide timely updates on children's progress to parents. Today KTQ has 15 Madrasas with a total of 1,500 students enrolled, involved in this initiative. I was awarded the Best Teacher award last year for my efforts in mobilizing and supporting the group. This was possible only because of ASER. (September 2, 2014)

[67] An ASER manager is a person hired for a temporary period to act as an extra ASER state team member during the first several phases of the ASER process, generally from training to recheck. ASER managers fulfill the duties of a state team member, but they generally have less experience with ASER. They can be thought of as existing somewhere between a state team member and a master trainer in terms of experience and capacity.

The "exposure" ASER afforded Fazal as he describes it sounds genuinely extraordinary—he leaves his hometown for the first time as a shy and self-conscious young man to conduct ASER and then, with the process completed, he returns having seen areas of his home state, Kashmir, for the first time and being utterly transformed in how he perceives himself. The "life lessons" sparked during his ASER participation burned brightly in Fazal as he returned to his community and his role as a teacher. Employing the skills he gained as an ASER Manager while utilizing the ASER findings as a starting point for collective deliberations about the status of education, Fazal engaged his community—fellow teachers, parents, friends, and others—in discussions that were instrumental to mobilizing them. They formed a grassroots group that took action to address the educational needs of their students and schools. Fazal not only credits ASER's enormous influence with changing the quality and trajectory of his professional life, but also with changing the educational opportunities for 1,500 students in his district. Although Fazal is likely an exceptional case in gauging what ASER can cultivate in its participants and staff members and the kind of lasting impact it can have on them personally and professionally, Fazal's resulting passion and capacity to bring about educational change cannot easily be dismissed. His story and the actions he undertook with others illustrate the vision that animates ASER and are a depiction of ASER's highest goals along with policy change.

Planning and Piloting for Upcoming ASER (Interim)

During the interim period between ASER wider dissemination and partner and volunteer recruitment mid-summer, the Centre begins ASER test creation (including the writing and piloting of samples) for the next cycle of ASER. Also, ASER Centre hires and orients new state and national staff members ahead of the recruitment and training period for ASER. This period is also a time in which ASER Centre can hold national capacity-building trainings for its more

veteran staff members, such as trainings on statistical analysis and presentation of data. Finally, this period is when the Centre engages in smaller-scale studies and projects that inform the NGO's understanding of educational issues in India and shape its hypotheses about why learning levels and other aspects of its findings are observed.

Discussion

In prior sections, I suggested that it is fruitful to analyze ASER as an evaluation, while also acknowledging that current evaluation theories have Western biases and limitations, which have two implications: First, biases hinder evaluation theories' current usefulness for designing evaluations in the Global South that can effectively address stakeholders' needs (Ofir & Kumar, 2013). Second, given their biases and limitations, evaluation theories are problematic as conceptual frameworks for research on evaluation—specifically in this study, for analyzing locally devised evaluations in India. As such, in the following sections, I candidly discuss ASER features in relation to evaluation norms or concepts but avoid fitting ASER neatly into how evaluation concepts have typically been defined. Wherever possible, I refrain from making judgments regarding ASER's worth or success according to how well ASER fulfills concepts or conventions. Toward utilizing my research on ASER to theorize about evaluation from the ground up, I aim to anatomize ASER's methodological features and process and then discuss the ways in which ASER may challenge the field's current ideas, so upon second glance ASER may enrich evaluation theory moving forward.

ASER as an Evaluation

Evaluation questions. If ASER is an evaluation, one of the most fundamental things to examine is, what questions guide its evaluation efforts? ASER has one explicit evaluation

question: "Are our children learning?"[68] This is the principal question that animated the design of ASER. Moreover, according to its architects, the outcome of learning is the most important indicator for judging the efficacy of India's primary education system (Goodnight, Paper 1). The "our" in the question is telling—it conveys how ASER seeks to promote a sense of shared responsibility for the quality of India's primary education system and for ensuring that each child is learning. It additionally connotes the collaborative nature of and the citizen engagement intrinsic to ASER.

There are complementary questions that are implicit in ASER's design, which further explain the reasoning behind the data that volunteers collect. They are the following:

1. How do children's learning levels relate to their enrollment status and grade level?
2. What is the status (i.e., quality) of rural government schools according to RTE indicators?
3. What household factors matter to children's schooling and learning?
4. What village factors influence children's schooling and learning?

Published annually in the tables of ASER reports are data related to ASER's principal question (i.e., "Are our children learning?") and the first two above-listed questions. Consequently, there is widespread debate in India about the credibility of ASER in answering these questions. However, the last two questions—though they evidentially motivate parts of ASER's design—remain largely unexplored in ASER reports.[69]

---

[68] This question was a recurrent theme in my interviews and field conversations with people participating in ASER—from ASER's architects to members of its partner organizations to its volunteers. When I asked "What is the purpose of ASER? What is it trying to do?" people responded that ASER is trying to find out and it is asking, "Are our children learning?" This question is now the recurring title of the educational reports published by UWEZO, a non-governmental organization that has adapted the ASER model and is conducting its version of ASER in Uganda and Kenya.

[69] Aspects of these questions are occasionally addressed in the three to four introductory studies and essays published at the front of each annual report. For example, Wilima Wadhwa in ASER 2014 utilized ASER household survey data (e.g., children's enrollment status, children's participation in private tuition, and families' affluence) to discuss patterns in rural children's participation in private schools and tutoring. She pointed to the significance of household factors in dispelling a widespread misconception in India that private schools consistently provide better

As previously introduced, an important query in designing and studying evaluations is, whose questions are being asked? And, relatedly, whose interests are being served? In the case of ASER, it is clear that ASER's architects—Drs. Chavan, Banerji, and Wadhwa—devised the questions being asked based on their shared perception of what was in the public's interest and what would benefit the government in improving primary education. Thus, there are inherently two audiences for ASER findings—ordinary citizens and government officials. The former category includes a breadth of stakeholders from parents to business owners that the architects identified as having innate investment in knowing the status of government schools and learning.

Citizen-engagement in the execution of ASER is intrinsic to ASER's design. However, the architects devised no mechanism for directly engaging ordinary citizens in the design process for ASER, specifically the determination of its evaluation questions. For some in the evaluation field, this may cast doubt on whose interests ASER serves because in evaluations that are intended to be "transformative," "democratic," and "equity-focused," one common concern is stakeholders' inclusion in the evaluation's design and their degree of influence over the questions the evaluation pursues (Greene, 2016; House & Howe, 2000; Mertens, 2009). Then again, the ASER architects did not perceive themselves as designing an evaluation—they were not operating with the field's concepts forefront in their minds. As such, the architects were not focused on adhering to evaluation conventions that guide how evaluators are supposed to facilitate stakeholders' self-determination or participation in the evaluation process. Instead, they generally saw their efforts as trying to prove a concept: simple-to-understand information about rural children's learning and government schools could routinely and reliably be collected and then used by the government and by the public to take action. What resulted is a form of

---

learning opportunities for children: Dr. Wadhwa unequivocally argued that comparing the learning levels of children attending government and private schools is not "comparing apples to apples" (ASER Centre, 2015, p. 19).

evaluation, but its achievement was organic and locally driven. Accordingly, the absence of citizen engagement in ASER's design process (and determination of evaluation questions) does not seem to be a fair proxy for understanding whether ASER is responsive to public interests or whether it adequately serves the public good. At the same time, in recollecting Ofir and Kumar's (2013) observation that responsiveness to stakeholder interests is a key ingredient in producing useful data, ASER's credibility and usefulness may be further enhanced if in the future, ASER can devise new opportunities for diverse stakeholders to share their questions and have them addressed via ASER's evaluation process and findings.

Audience for findings. Regarding the intended audience for ASER's findings, its annual report has largely been aimed at sharing data with Indian officials in different layers of the government (i.e., at national, state, and district levels). A main goal of ASER from its creation was to pressure bureaucrats and politicians to use evidence in their decision-making. This is one reason why indicators of school quality according to the Right to Education Act's benchmarks are presented in report tables alongside learning levels (see ASER Centre, 2015 and Goodnight, Paper 1). As ASER's reputation has grown in the Indian media—including in national and state, and Hindi-, English-, and regional-language outlets—so has ASER's audience grown to include the educated public if not the wider Indian populace.[70] Increased media coverage has also contributed to popular perceptions of the Centre as a government watchdog and of ASER as promoting accountability. Meanwhile, the domestic and international audiences for ASER continue to grow as it is noticed by people in civil society organizations, foundations, and

[70] ASER's growing coverage in the media is not happenstance. The Centre has a communications unit dedicated to disseminating ASER findings and the directors are savvy in promoting the organization's work. Thus, success in raising ASER's profile is reflective of sustained efforts to generate interest in ASER and awareness regarding the issues with India's primary education system and children's learning.

academia; the innovativeness of ASER's model has meant that its results as well as its design have substantial influence in such arenas (Goodnight, Paper 3).

Despite all of these developments in the reach of ASER's findings, a considerable limitation of ASER is the extent to which its data are returned to the communities from which they were collected. ASER Centre has not yet discovered how to systematically and sustainably reengage surveyed villages due to the scale of ASER and the expert skills necessary for competently analyzing data and clearly explaining the meaning of data to rural stakeholders. The resources, time, and capacity (i.e., expertise) required for deliberating ASER results with communities appear insurmountable challenges within ASER's current structure and process, yet the Centre has been experimenting with different approaches like soliciting and funding partner organizations to disseminate ASER findings in some villages. This work and challenge continues.

Evaluator. In thinking about the ways that ASER does and does not seem to fit conventional conceptions of evaluation, one of the most interesting issues to consider is, who is the evaluator in ASER? Because evaluations are often viewed as "arguments in which evaluators [present] evidence for and against" (House, 2013, p. 200), the integrity and competence of an evaluator is a recurrent theme in the field's literature. Marvin Alkin (2013) defines the evaluator as a "specialist" in:

1. "group processes who helps potential evaluation users identify important aspects of their program,"
2. "examining the many aspects of an evaluation context,"
3. "identifying a broad array of potential means of acquiring information about programs and in implementing systems to obtain those data,"

4. "analyzing a broad array of types of information,"
5. helping "users themselves specify the means of valuing and identifying the potential action implications of valued findings" (pp. 287-288).

The qualities of an evaluator as defined by Alkin are somewhat difficult to reconcile within ASER's model. Perhaps, this is because the "specialist" role of an evaluator is actually distributed across many individuals in ASER. For instance, if one aspect of an evaluator's work is to design the evaluation's methodology and data collection instruments; this design work was initially undertaken by ASER's architects but is now shared by the ASER Centre Assessment and ASER Survey units in Delhi along with ASER state teams playing an instrumental part in developing and piloting ASER language tests. Another potential role of an evaluator is capacity-building—in ASER, this is a function spread across more than a thousand trainers from the central and state team members to the master trainers and partner organization coordinators located in the districts of each state. These trainers are who build the ability of others to examine the "many aspects" of ASER's "evaluation context" as it pertains to a particular scope of responsibility—that is, a state, district, or village. Finally, evaluators typically fulfill an important role concerning data collection and communication with stakeholders. In the case of ASER, interfacing with stakeholders on the ground and collecting survey and test data from them is work done by thousands of volunteers, who share in that most primary of evaluator functions as do their ASER master trainers, partners, and ASER central and state teams.

Evaluation literature routinely comprises meditations on the essential qualities of an effective and just evaluator—their expertise, ethics, judgment, interpersonal skills, and cultural competence (Alkin, 2013; Chouinard & Cousins, 2009). A range of qualities and capacities manifestly matter within the ASER model too, yet given that numerous participants appear to be

fulfilling different aspects of the evaluator's roles, various qualities essential to initiating, facilitating, and safeguarding a credible and effective cycle of ASER are in turn required from a diversity of individuals. In short, to speak of "the evaluator" in ASER is in some sense to speak of the requisite qualities of more than 25,000 people (annually). As importantly, Alkin (2013) argues the "evaluator is a change agent" who helps people "to see opportunities for improving their programs through the use of evaluation information" (p. 288). One innovation of ASER is in how it blurs the line between evaluator and stakeholder—or, between specialists and "ordinary citizens learning to measure what affects their lives." One can conceive how nearly any participant in ASER from staff member to volunteer has the potential to become an agent of change, yet it remains difficult to gauge how regularly ASER enables individuals to manifest change and how sustainable those changes are.

Wrestling with Evaluation Concepts in Analyzing ASER

Data credibility and usefulness. If evaluation is "determining" the "worth" of something (Scriven, 2013, p. 170), the credibility of an evaluation's data is a significant issue. One question perennially important within evaluation is, how *credible* are data for judging the condition, quality, or equity of the evaluand? A complementary question raised by Ofir and Kumar (2013) is, how *useful* are data? Both the credibility and usefulness of data are essential in determining whether data can provide "actionable evidence" for enhancing decision-making and practice in order to improve the worth of a program, system, or policy (Donaldson, Christie & Mark, 2015). Nevertheless, like determining the worth of something, the credibility and usefulness of evaluation data reflect subjective determinations, which can vary according to the particular experiences or interests of a given stakeholder.

ASER architects were concerned about the usefulness of ASER data during the evaluation's initial design (Goodnight, Paper 1). While the type of information collected was aimed at being useful to the planning and budgetary arms of the Government of India, ASER Centre also endeavored to show state governments, the public, and the Ministry of Human Resource Development (i.e., the central government body directly charged with the provision of education) India's degree of progress with universalizing education across rural districts. Interestingly, ASER indicators were not crafted to produce data specifically useful to educators—ASER data are not suitable for planning curriculum or for reforming instructional practices. In short, ASER data are oriented towards political and managerial functions versus serving the needs of education practitioners.

ASER Centre and its partners care about the extent to which ASER produces "actionable data"—the simplicity of ASER data and the starkness of ASER's results, published year after year, are meant to spur popular action. Though ASER findings were initially aimed at those in power with decision-making and managerial responsibilities, the concurrent emphasis on transparency and integrity in the conduct of ASER and on open dissemination of ASER protocols and results reflects the architects' desire for public use of ASER data (Goodnight, Paper 1). ASER Centre directors expressed strong committed to maximizing the meaning of ASER data for multiple users. From the Centre's perspective, the more partner organizations and everyday citizens who access ASER findings, the more people who are engaged in deliberations about the status of children's learning and primary education. Obviously, ASER data use currently requires from intended users a minimum threshold of educational capacity and ability to access findings (e.g., via the internet), which is still beyond the capacity and ability of many villagers from whom volunteers collect ASER data.

Because the focus of my paper has been ASER's evaluation design and process versus its long-term impact, my discussion regarding the usefulness of ASER's data is limited. Given this study's scope, it is easiest to examine ASER's goals for data use versus its success in this regard. Similarly, this study is limited in its assessment of the credibility of ASER data from the perspective of different stakeholders in India. Nonetheless, it is possible to identify the strategies the ASER model employs to guarantee the trustworthiness of its data. In this regard, this papers' sections on ASER monitoring and recheck demonstrate that ASER Centre is dedicated to safeguarding the integrity of its findings. Continued improvements in this area exhibit that the Centre is tremendously committed to data quality. Despite ASER's tremendous scale and that volunteers collect its data, the proliferation of its monitoring and recheck efforts enables the Centre's leadership to mount a convincing case for the credibility of ASER data.

Social accountability. As opposed to a narrower conception of accountability, Christie and Alkin (2013) posit that "true accountability requires 'answerability'—that is, those responsible must be held accountable" for the quality and equity of social programs (Christie & Alkin, 2013, p. 14). They claim that evaluation essentially "provides the information for 'being answerable'" (p. 14). Within an open democratic society, social accountability is a crucial value—that is, citizens trust that elected officials fairly administer public resources and competently manage public systems to serve the common good. Maintaining public trust is deeply related to the legitimacy of evaluation: Social accountability "legitimizes evaluation as a fundamental process for generating systematic information for decision making…. within the configuration of government-sponsored programs and policies" (Christie & Alkin, 2013, p. 15).

The theme of social accountability is integral to the ASER model—apparent within its philosophy, design, and process. The leaders of ASER Centre actively promote the need for

government answerability in the provision of primary education and in effective use of public monies. They assert the government should be answerable from the very top downward to state, district, and village bodies. However, a key aspect of understanding the culture of ASER is recognizing that this concept of answerability also applies to ASER staff members, partner organizations, and volunteers. ASER fundamentally expands who is accountable or answerable for the right to education in India to all "ordinary citizens." Thus, the ASER vision of "ordinary people learning to measure what affects their lives" is a sociopolitical statement that social accountability cannot be disentangled from individual accountability in a democracy.

ASER advocates a view of citizen rights that is heavily reliant on a concurrent recognition of personal responsibility to the collective good. Everyday people are imbued with the right to have, but also the responsibility to ensure, good education. A manifestation of this philosophy is the role that integrity plays in the culture of ASER. Sincerity is considered essential in facilitating the ASER evaluation. In guaranteeing the ultimate quality and credibility of ASER, the honesty of ASER volunteers, MTs, and state and central team staff members is so fundamental and their willingness to be transparent about issues they encounter during the ASER process is so crucial that it makes it undoubtedly clear why the ASER pledge has so heavily emphasized this theme of personal responsibility. Year after year, the pledge is the culminating moment of every training when "ordinary citizens" officially become ASER participants—whether volunteers, master trainers, or state team or central team members, they all pledge their "utmost sincerity" and "complete honesty" in protecting the "vision of those who have come before." Furthermore, they solemnly pledge through their work in ASER to "become an example for those who will follow…in the years to come" (Appendix E).

In the evaluation and reform of social systems, ASER's apparent blurring of the lines between government and citizenry can be interpreted in multiple ways. From one perspective, ASER Centre's vision seems to diminish the government's direct responsibility for providing free primary education and for ensuring its quality. ASER advocates that it is also citizens' responsibility to independently monitor schools, to consume evaluation information, and to ensure equitable, high-quality schooling. This concept of citizens challenging the government but also working alongside the government and helping to supplement the efforts of government wherever it falls short is a value central to Pratham that has defined Pratham's grassroots educational programming (Goodnight, Paper 1). Accordingly, ASER provides information not just for government but also for civil society; the Centre believes all individuals must be willing to take action to improve social systems. From another perspective, ASER is fiercely democratic. Through their ASER participation, people are encouraged to develop their own capacities and sense of agency. ASER Centre wishes for its staff, partners, and volunteers to be motivated, self-determined, and deliberative in their appraisal of information, in their organizing efforts, and in their actions.

Evaluative thinking. Evaluations should encompass opportunities for "cultivating the critical voice of others" (Schwandt & Gates, 2016, p. 68). In other words, evaluation processes should strengthen the capacity of stakeholders to deliberate and think "evaluatively" about their social programs or social systems, as Jennifer Green (2016) states: "To think evaluatively is to question the logic and practical sensibility of the program, to wonder about how people are experiencing it, and to look for evidence on how well the program is making a consequential difference in people's lives" (p. 49). She reveals that to "think evaluatively" is not an apolitical exercise in critical thinking, but it has a strong equity component because it is also "to ask how

well the program's resources are reaching those least well served, [and] to ask if the program is affording equity in access, experience, and accomplishments to all" (p. 49).

The training, data collection, and recheck phases of the ASER process provide several opportunities for ASER participants to grow in their ability to think evaluatively about how India's primary education system is "making a consequential difference" in rural "people's lives." For instance, after receiving training on ASER's goals and how to execute it, volunteers while collecting data in villages can ask how well this system is "reaching those least well served" based on their firsthand observations of the government schools and children's learning levels. Later, volunteers can reflect on the rural households they visited and the conversations they had with families to determine independently or in discussion with one another if the district's school system is "affording equity in access, experience, and accomplishments to all." What can be ascertained from ASER's process are there appear to be ample chances for developing one's capacity to think evaluatively and what is clear from ASER's vision statement is that this an explicit goal of the evaluation.

The extent to which individuals actually do develop in their ability to think evaluatively resulting is not something this study can directly answer. Rather, this investigation of the ASER model can highlight the potential for evaluative thinking in the ASER process. Regarding rural families and communities, the process only allows for their brief encounter with volunteers during data collection—while this process results in some meaningful conversations about education and important revelations about their children's basic learning, ASER does not allow for sustained engagement with villagers. Thus, the ASER process is not affording rural communities much opportunity to develop their critical voice or ability to think evaluatively about education.

The potential for long-term benefits regarding evaluative thinking and cultivating a critical voice is greater for ASER staff, MTs, partner organizations, and volunteers. Even so, there is logically high variability between and within these different groups as the capacity to think evaluatively and to speak critically depends upon type of participation as well as individual characteristics like initiative. Fazal's blog post suggests the profound success ASER has in cultivating these abilities in some people with key factors being the length of one's experience (i.e., number of rounds of ASER) and the quality of one's participation in ASER. The many unsettled questions regarding ASER's influence on persons' capacity to think evaluatively comprise fertile ground for future research.

Social conscience. It is helpful to revisit this notion of evaluation as a form of "social conscience" that raises a moral voice concerning important social, political, and cultural issues (Schwandt & Gates, 2016, p. 67). According to the description offered by Schwandt and Gates, asking whether ASER itself acts as a kind of social conscience entails examining whether it engages "significant inequalities" and "power differentials" in the access to and provision of social systems (Schwandt & Gates, 2016, p. 67). Further, such an analysis prompts an inquiry into whether ASER asks "difficult questions about what kind of society we should have, what directions we should take" (Schwandt & Gates, 2016, p. 67). Finally, we can investigate if ASER is a "risky" enough endeavor, which raises "serious questions of public direction" and dares to explore "what we do not know about what we are doing" (Schwandt & Gates, 2016, p. 67).

By examining education in the historically underserved rural regions of India, ASER is indeed a kind of social conscience. While providing its participants with life-changing exposure to rural schools and households, the sections preceding this one illustrate how the ASER process is intended to collect data about the school system in a way that facilitates deliberation about

"what direction" the government and citizens "should take" to improve it—ASER advocates collective responsibility and collective action as detailed in the social accountability section.

A notable limitation of ASER is its lack of data on tribal, caste, and religious identity, which influences the potential of ASER results for challenging certain types of power differentials in the education system. ASER Centre's inability (or refusal) to collect these types of household data prevents the disaggregation of learning levels according to significant social identity markers. The continued importance of these identities as sources of acute inequality in Indian society is clear, and they are consequential factors in educational access, experiences, and attainment (see Goodnight, 2017).

This issue was raised in interviews with different members of ASER Centre's leadership, which resulted in a couple explanations given for why children's tribal, caste, and religious identities are not collected during the household survey. One explanation offered is that it is extremely difficult to collect reliable data on these identities from families because there are countless caste-based groups across regions and a plurality of names may be used to describe them. Furthermore, household respondents may be incentivized to lie about their identity if they believe their responses determine eligibility for government benefits. Another explanation provided was that the Centre wants ASER data to be apolitical. If the Centre collects data showing disparities in educational provision, access, learning or in household affluence along caste, tribal, and religious lines, ASER data can be used by various political groups to mount arguments for purposes other than that for which ASER data are collected. The Centre does not want to be associated with these arguments or to be seen as feeding information to identity-based political groups. Evidently, in order for ASER Centre to retain the support of its various partner organizations and continued cooperation from state governments and DIETs, the organization

needs to avoid courting this type of controversy. As the previous sections illustrate, such cooperation and partnerships are fundamental to the execution of ASER.

Regarding the second explanation, I do not interpret ASER as existing in an apolitical space. ASER's goals, process, and findings are political regardless of the Centre's stance on this specific issue. The complementary explanation appears to make more sense: ASER data are currently not of interest to some political groups whose core mission is to agitate for the power or rights of specific caste or religious groups, but this could change with the collection of new social identity information and the Centre sees it in ASER's best interest to maintain the status quo. Interestingly, ASER data are disaggregated by gender, so the Centre is not avoiding the presentation of data that highlights discrimination in other regards. Barring problems with reliability, why should gender be seen as distinctly different or a more important issue than caste, religion, or tribal identity in an analysis of education? Does understanding the learning outcomes and educational participation of children from these groups not matter as a key issue of education justice? It is difficult for me to accept ASER Centre's stance on this point. However, I appreciate that the ramifications for the organization in undertaking such analyses are different. Though, exploring gender discrimination can also be a controversial undertaking.

The disaggregation of education data to enable analyses by different social identities and for marginalized groups is a worldwide justice issue, according to Pauline Rose (2015), former Director for UNESCO's Global Education Monitoring Report, because such analyses are significant to every nation's understanding of their true progress in universalizing educational opportunity. Education data sets must represent all types of children and allow for researchers to identify them by collecting relevant social identity indicators. In terms of representativeness, ASER succeeds where other data sets fail, including a wide range of children by sampling from

households instead of schools (Goodnight & Bobde, unpublished). Because of ASER's scale and the overall trustworthiness of its education data, it is poised to have a monumental impact on the study of educational inequity in India. The ASER methodological design and process already reflects considerable ingenuity in how it overcomes commonplace barriers to collecting education data. ASER Centre is an organization with inordinate educational and social research experience in India that is capable of experimenting with and piloting different approaches for how to reliably collect information on caste, tribe, and religion from households, if it so chooses.

Experimentation in the collection of social identity information in ASER would have to rely heavily on the local knowledge of ASER state teams and partner organizations who are likely best positioned to make judgments about how to ask individuals in households for such information. Most state teams and partners are probably already aware of what information is consequential in school participation and learning outcomes in their areas as repeated participation in ASER facilitates the acquisition of incredible qualitative understanding of the contexts in which one works. Based on my three-state study of ASER data collection, I gather the same identity markers do not matter equally to educational opportunity countrywide: there are important regional differences. For instance, I have been told that caste and gender identity in Manipur are less significant to educational participation than tribal identity and mother tongue. Apparently, this is not the case in Rajasthan where all four factors are significant. In Tamil Nadu, caste and gender seem important factors in educational decision-making and outcomes but perhaps to a different degree than in Rajasthan. The ASER process probably can accommodate some experimentation with different strategies for collecting additional social identity data as the evaluation has trialed additional tests and indicators in the past. ASER Centre may not publicly report such data until the organization is confident that it is reliable.

The reality for ASER is that without its continual willingness to risk anew its role as a form of social conscience may fade. Specifically, I believe ASER should risk beginning and failing in the collection of data for these crucial markers of social identity—caste, religion, and tribe. Otherwise, important strides forward in the relevance or usefulness of its data may go unmade. Over a decade ago, ASER was born out of a daring spirit. Today, it still reflects a strong moral vision for what evaluations can aspire to be. ASER offers hopefulness and a direction for collective action in situations where the barriers to learning seem many and the challenges for policy and the school system seem overwhelming. ASER is where the cold technical practice of evaluation meets the warmth of humanitarianism and the messiness of divergent lived realities.

Conclusion

> *"After the formalities of training and survey get over in the evening, ASER seems more than just a survey. It has been an opportunity to overcome our fears and prejudices about people and place, to learn to strike friendships at out of the ordinary places, to get a glimpse of children's lives…. The remarkable thing about ASER is that as it has grown, it has helped volunteers like me grow alongside."* –Saachi, blog post entitled, "Growing up with ASER" (October 14, 2014).

During the methodological debates of the 1970s where evaluation experts and theorists disagreed about the merits of quantitative versus qualitative approaches, Ernest House (2013) recalls that he asserted evaluators "might use both quantitative and qualitative data" (p. 200). House (2013) stated that at the time he "looked for a broader perspective" and subsequently conceived evaluations as "arguments in which evaluators presented evidence for and against"; thus, in his view, "evaluation was more than methods" (p. 200). In fact, evaluation was "more than literal truth"—it was "highly political" and ultimately about "justice" (pp. 199-200).[71] I am similarly advocating for "a broader perspective" of monitoring and evaluation that includes the

[71] House claimed: "Evaluations should be true, coherent, and just"—if evaluations are inadequate in any of the three, they are "invalid" (House, 2013, p. 200).

innovations and methods adopted by M&E efforts aimed at large-scale social systems. Throughout this paper, I have argued that the inclusion of ASER as a model will allow for better theorizing on how large-scale evaluations are about more than the "literal truth" of numbers and relate to social interaction, to politics, and to issues of justice.

The academic literature on evaluation can beneficially inform activities that have been traditionally labeled "monitoring." The purposes of evaluation versus monitoring are increasingly blurred and outcome indicators are used more often to inform improvement in large social systems like education. If evaluation theorists study endeavors like ASER, given that such M&E efforts are becoming commonplace in valuing social systems and public policies (see Gildemyn, 2014; Goodnight, Paper 3), I believe the strength and breadth of evaluation theories will subsequently be bolstered and in turn, some typical problems with M&E may be ameliorated. Also, this paper demonstrates that analyzing ASER in light of the conceptual, ethical, and practice considerations raised by evaluation theories is fruitful.

In the case of ASER, I have argued that the dichotomization of monitoring and evaluation is especially problematic. Because of the negativity and trepidation sometimes ascribed to monitoring both within and beyond the field of evaluation, there is the risk of arbitrarily dismissing the complexity of ASER's process and the ingenuity of its design. If ASER is dismissed as monitoring, its social meaning and its political effects are in danger of being overlooked. For instance, conventional views of monitoring cannot account for the lengths at which ASER participants go to stimulate the collective social conscience of India.

ASER is a groundbreaking citizen-engaged model for the evaluation of social systems: The ASER evaluation model is a synthesis of components of program evaluation, policy evaluation, and monitoring (see Goodnight, Paper 3). M&E efforts increasingly tackle giant

social systems and are charged with producing system-wide data for public consumption. By including models like ASER in mainstream theorizing on evaluation, the field can strengthen the ethics, responsiveness, cultural competence, and overall efficacy of future monitoring and evaluation efforts designed for large social systems. Also, the ASER model pushes evaluation theory into directions that serve current needs to provide system-level data or information that addresses the pressures on decision-makers at high-levels of governance who struggle with how to use data effectively and make "actionable" the findings that are yielded from conventional evaluations. By applying evaluation theory to M&E efforts that are aimed at judging the quality or equity of large social systems, perhaps the field can make monitoring more self-reflective. Also, such theorizing can bring other researchers and practitioners like economists, statisticians, sociologists, development specialists, and educators, who use big data, into the fold of the field of evaluation and its academic literature. This means creating a literature that addresses these researchers and practitioners where they are currently at in terms of their understanding and reflectivity regarding the use of large-scale data for evidence-based decision-making.

The annual report published by ASER Centre—thick with multicolored tables and graphs that reveal ASER's findings—may lead one to think that ASER is principally technical in nature. Furthermore, the protocols and quality control mechanisms of ASER corroborate perceptions that evaluation is principally "the systematic collection of information" (Patton, 1982, p. 35). However, I argue that ASER is as much a social process as a technical one. In fact, people who participate in ASER and the ASER pledge itself call it a "movement"—that is, a collective of citizens working together to advance a social and political cause. One of ASER's three architects, Dr. Madhav Chavan, characterizes ASER's creation as a decision "in favor of rebellion" (Goodnight, Paper 1). His statement reveals that the spirit underlying ASER is one of

trying to bring about a new social order in India wherein people recognize their right as citizens within a democracy to independently evaluate the efficacy of public institutions (i.e., government schools) and social systems (i.e., primary education system). Nonetheless, with that right, ASER asserts that citizens have a complementary obligation—they are collectively responsible for action to improve their social systems and, perhaps as Saachi suggests, to challenge their own "fears and prejudices about people and place" that may prevent social systems like education from serving all children everywhere well.

PAPER 3. *Design Influence*: The Source of an Evaluation's Power Beyond Process and Results

Abstract: The *Annual Status of Education Report (ASER)* is produced via a pioneering large-scale evaluation of primary education in India. The process underlying ASER, which includes village mapping, school surveys, household surveys, and household-based learning tests, relies on the participation of thousands of volunteers and hundreds of partner organizations every year. Based on a 10-month ethnographic study of ASER, this paper addresses how ASER reflects *design influence*, that is the influence of an evaluation's entire concept—from its goals, to its capacity-building components, participatory structure, data collection instruments, and dissemination strategies. Evaluation designs like ASER can ultimately influence individuals and institutions not involved in the initial evaluation because of the design's innovativeness in addressing planning needs. These designs are elevated as models. Within the international "policy-shaping community" (Cronbach, 1982), ASER's approach has become a solution for closing the global data gap on learning and primary education universalization.

Each year, tens of thousands of volunteers for the *Annual Status of Education Report* (ASER) collect data on foot through summer heat and monsoon rain in villages spanning vast India. ASER volunteers evaluate the conditions of government primary schools, record the enrollment status of school-age children, and assess their basic learning levels in reading, math, and English language. With the help of approximately 500 partner organizations and 25,000 volunteers countrywide, the Delhi-based nongovernmental organization (NGO) ASER Centre conducts ASER in all of India's rural districts, independent of the Government of India (GOI).[72] ASER's final product is a thick, widely publicized report. To generate the report's data for each district, a couple of ASER volunteers travel to a sampled village and gain support from the village head before proceeding with data collection. The data collection process for ASER includes mapping and surveying the village, surveying the local government primary school, and systematically selecting and surveying 20 village households. During household surveys,

[72] ASER has been conducted every year since 2005, with the exception of 2015 when it was suspended to allow ASER Centre leadership and staff to pursue pilot research related to expanding what ASER currently measures through its assessments and surveys.

volunteers also administer basic reading and math tests orally and one-on-one to all children (ages 5-16 years) living in the home. This process is repeated for 30 villages for each district in every state of India.

The purpose of ASER's report is to present a real-time evaluation of rural primary education in order to prompt public dialogue about the status of children's learning, the progress of India's universalization efforts, and the condition of its educational governance (Goodnight, Paper 1). ASER offers the only annual and publicly available figures on children's learning levels pan-India. The report's 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 language skills (see ASER Centre, 2015).[73] Several design innovations not only enhance the representativeness and validity of ASER findings, but they make its massive scope possible (Goodnight, Paper 1): in ASER 2014, volunteers surveyed 341,070 rural households in 16,497 villages across 577 districts of India and administered learning tests to 407,706 children (ASER Centre, 2015, p. 316). Importantly, ASER findings include Indian children who are typically missing from large-scale educational data, such as children who attend unrecognized private schools or who have dropped out of school, as a result of one of ASER's crucial methodological innovations: its tests are administered in rural households, not schools (Goodnight & Bobde, unpublished). Moreover, as compared with commonplace 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

[73] ASER is designed to produce representative estimates at the district-level. In 2009, Ramaswami and Wadhwa conducted a study that looked at the precision of ASER state and district level estimates, which found that whereas state level averages were within "a margin of error of 5% or less," district level estimates were "less precisely estimated" because ASER's second stage of clustering (i.e., household selection) "increases the variability of estimates" (ASER Centre, 2015, p. 275). Thus, ASER 2014 now reports divisional estimates for districts from most states, but not for some smaller states with few districts (and smaller populations of children within districts) like Manipur.

limited literacy and exposure to schooling (and thus, who would not be able to easily read and follow directions for completing test tasks independently).

The design and process underlying ASER present a unique model for the evaluation of large social systems, like education (Goodnight, Paper 2).[74] The ingenuity of ASER has evidently led to the broad influence of not only its findings, but also its design, which has gained the attention of a globalized "policy-shaping community" (Cronbach, 1982), which operates within and beyond India's borders (e.g., Barrett, 2011; Jha & Parvati, 2014; Learning Metrics Task Force, 2013; Nawani, 2013; Pritchett, 2013).[75] Since 2008, ASER's model of evaluation has also proliferated across countries in the Global South largely due to its cost-effectiveness, household-based approach, capacity-building components, and inclusion of volunteers and partner institutions, all of which allow for data to be collected by civil society organizations and independent from governments (Results for Development Institute, 2015).[76] To date, ASER's *citizen-engaged evaluation* model has been adapted by NGOs and research organizations across three continents for use in thirteen countries.[77]

[74] Some may view ASER's model as "performance monitoring" versus evaluation (see Chouinard, 2013). I argue monitoring is a form of evaluation that should draw upon (and inform) evaluation theories, ethics, approaches, and methods and furthermore, that ASER is not a conventional monitoring exercise (Goodnight, Paper 2).

[75] Lee Cronbach's "expansive concept of the policy-shaping community" envisioned an audience for evaluation, which included "policymakers, civil servants, interested citizens, advocacy groups, and the media" (Greene, 2013, p. 98). In the case of ASER, its globalized audience includes multilateral and bilateral policymakers, advocacy groups, and international and national media who are focused on education and development issues.

[76] When I refer to ASER's evaluation "model," it comprises ASER's methodological design and entire process for training, data collection, and monitoring as well as the data subsequently produced from its surveys and tests and the report itself. Also, included in this model is the Centre's process for the dissemination of ASER findings.

[77] The People's Action for Learning (PAL) Network formed to support cross-country efforts to employ the ASER evaluation model (called "citizen-led assessments" by the PAL Network), with financial support from the Hewlett Foundation. According to the PAL Network website (http://palnetwork.org/), the following organizations in the following countries are involved: MIA in Mexico, Jàngandoo in Senegal, Bɛɛkunko in Mali, SCALE in Ghana, LEARNigeria in Nigeria, Djangirde in Cameroon, Uwezo in Uganda, Uwezo in Kenya, Uwezo in Tanzania, TPC in Mozambique, IID/BRAC Survey in Bangladesh, ASER in Pakistan, and ASER in India.

This paper explores the reasons for ASER's influence, specifically the qualities of ASER's evaluation design and process that explain the global momentum behind the spread of its model. The paper includes six sections: following this introduction, the next section is a review of literature on *evaluation influence*. Afterward, I outline my study's questions, its methods, and data. The fourth section offers a brief history of the development of ASER's design and process as well as explains ASER's philosophical underpinnings. In the discussion section, I examine ASER according to the various sources of its influence. The paper concludes with an explanation of why ASER's model has become globally influential.

Theoretical Framework

> *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)

Karen Kirkhart (2000) has written about *evaluation influence* as building upon the notion of *evaluation use* to conceptualize an integrated, fuller understanding of evaluation's meaning.[78] Evaluation influence importantly allows for the consideration of evaluation effects that are "multidirectional, incremental, unintentional, and noninstrumental alongside those" effects that have been the conventional aims of evaluation practice and are more generally associated with evaluation use (i.e., "unidirectional, episodic, intended, and instrumental" effects). Though not without its critics (Alkin & Taut, 2003; Nunneley, King, Johnson, & Pejsa, 2015), evaluation influence as a theoretical framework enables researchers of evaluation to cultivate a more holistic conceptualization of evaluation in investigating its implications for stakeholders and the *evaluand* (i.e., the program, policy, or social system under evaluation). Evaluation influence has

[78] Alkin and Taut (2003) define *evaluation use*, or "evaluation utilization," as "the way in which an evaluation and information from the evaluation impacts the program that is being evaluated" (p. 1).

not been as fully articulated as the many facets of evaluation use (Nunnelley et al., 2015) like *instrumental use*, *conceptual use*, *symbolic use*, and *process use*.[79] Nevertheless, the broader notion of influence offers a framework for theorizing an evaluation's "immediate," "end-of-cycle," and "long-term" consequences as a manifestation of the evaluation's process or results (Kirkhart, 2000, p. 8).

Kirkhart's (2000) integrated theory of influence," illustrated in her diagram (Figure 13), highlights the relationship between three dimensions of evaluation influence:

- **Source:** The first dimension is the "source of influence," which 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 and "the data-based conclusions" that can be drawn from it (Kirkhart, 2000, pp. 8-10). Based on the source of influence, two kinds of evaluation influence, *process-based influence* and *results-based influence*, are broadly conceived.
- **Intention:** The second dimension of intention describes "the extent to which evaluation influence is purposefully directed, consciously recognized, and planfully anticipated" (Kirkhart, 2000, p. 11). This second dimension examines whether the effects of an evaluation's process or results are intended or unintended according to the evaluation's purpose, goals, and operating theories (i.e., philosophy) that are embedded in the evaluation's design. Three aspects of intention are identified: 1) the type of influence

[79] Instrumental use refers to situations where the findings from an evaluation are employed in taking "direct, visible action" whereas conceptual use is when evaluation results lead to "enlightenment" or "demystification," but they do not "necessarily lead directly to change in overt behavior" (Kirkhart, 2000, p. 9). Symbolic use—sometimes called political use or legitimative use—is "where evaluation was used to either justify a decision that had been previously made or to demonstrate that a program was willing to be evaluated, thus enhancing the reputation of the program manager or decision-maker" (Alkin & Taut, 2003, p. 5). Symbolic use is about using evaluation or evaluation findings because of what they signify and for principally political or appearance purposes. Process use is "not the use of the evaluation findings [i.e., results], but referring to the manner in which the conduct of the evaluation (the evaluation process) impacts on individuals or organizations" (Alkin & Taut, 2003, p. 5).

desired or anticipated, 2) who is to be influenced by the evaluation, and 3) the processes, findings, or person (e.g., evaluator, program manager, and policymaker) that/who are expected to exert influence.

- **Time:** The third dimension pertains to "time frame" and the "chronological or developmental periods in which evaluation influence emerges, exists, or continues" (Kirkhart, 2000, p. 9). The sub-divisions of the time dimension identify "the need to recognize influence during and immediately following the evaluation cycle as well as effects that are visible in the future" (p. 9).

Kirkhart (2000) cautions that the subdivisions represented in her diagram are "admittedly somewhat arbitrary" and "may be more accurately characterized as continua" because the realities of an evaluation's influence are in the "gray areas that fall between" (p. 8). Thus, while providing the field with a theory for the interpretation of evaluation influence, Kirkhart (2000) also cautions against the use of its dimensions toward dichotomous thinking and forming rigid categorizations. Additionally, it is possible that a particular evaluation may have relatively little influence or contrastingly, multiple influences—some intended and some not—emanating from different sources and occurring at various times.

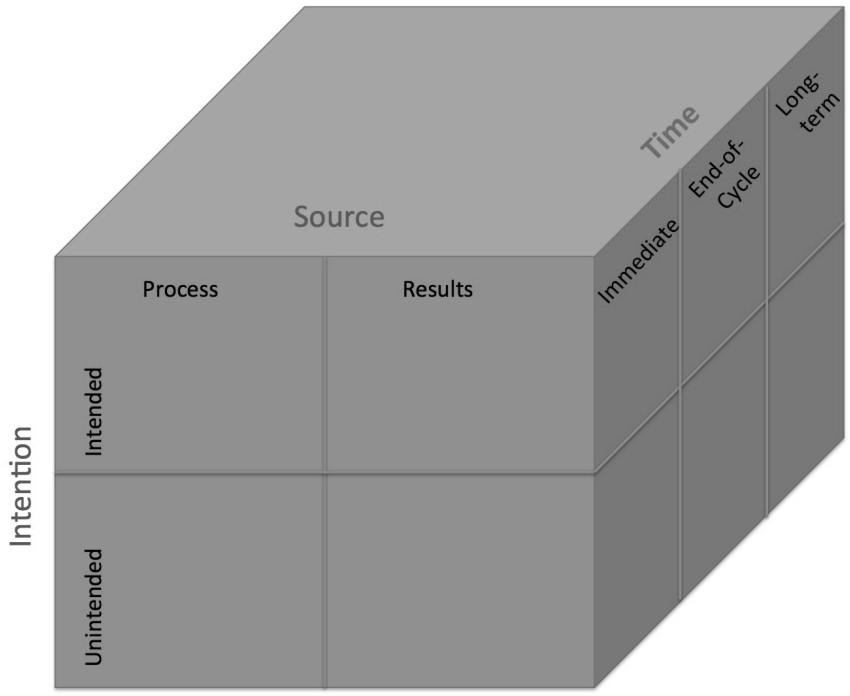

Figure 13. Kirkhart's (2000) diagram for her integrated theory of influence. Adapted from "Reconceptualizing Evaluation Use: An Integrated Theory of Influence," by K. Kirkhart, 2000, *New Directions for Evaluation, 88,* p. 8. Copyright 2000 by Jossey-Bass Publishers. Adapted with permission.

Social Betterment and Levels of Influence

Henry and Mark (2003) expand upon Kirkhart's conception of evaluation's influence, suggesting that influence is "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" (p. 293). Thus, Henry and Mark (2003) envision evaluation influence as dynamic processes that inspire consequential, interim changes toward a long-term goal: social betterment. They define social betterment in a "democratic society" as "bringing about a state [in social conditions] that would be considered as better than the state that existed before, as judged through deliberation and by public opinion" (p. 295). Furthermore, Henry and Mark suggest a broad base of stakeholders (i.e., the public) as the judges of whether evaluation activities contribute to social betterment, and they propose deliberation as the proper means for making such a judgment.

Beyond their articulation of social betterment, another one of Henry and Mark's (2003) contributions to theorizing on evaluation influence lie in their focus on the interim outcomes of an evaluation that shape its "influence pathway"; resulting from an evaluation's process or results, they identify these interim outcomes (i.e., change processes) occurring at three levels: the individual, interpersonal, and collective. Individual-level influences are characterized as instances when the evaluation motivates changes ("in thoughts or actions") predominantly "within the individual" (Henry & Mark, 2003, p. 297). Interpersonal-level influences refer to changes to a "process or outcome that predominantly takes place within interactions" between people (p. 298). Collective-level influences are changes that are institutionalized in that they happen within public or private organizations; in other words, the change happens to a process or outcome that "operates predominantly within some aggregate social organizational body" (p. 298). While concurring with Kirkhart's conceptualization of evaluation and acknowledging the dimensions that she identifies, Henry and Mark (2003) seek to infuse evaluation with the "moral compass" of social betterment and propose that evaluation influence emanates from interim changes at the individual, interpersonal, and collective levels (p. 295).

Evaluation Language: Disentangling Influence from Use

Language in evaluation, as in other fields, is important because it shapes understanding and creates a foundation for the exchange of ideas. Given the evaluation field's interdisciplinary and organic development, there are unsurprising issues with the use of terminology (as Kirkhart suggests), which expose conceptual misalignments in its literature (Goodnight, Paper 2). According to Kirkhart (2000), the "scope and language of past approaches" have constrained appraisals of evaluation's influence and full meaning (p. 5). Kirkhart's (2000) proposal of the term *evaluation influence* is an attempt to overcome the narrowness and "fragmentation" of past

conceptualizations of evaluation's effects that have been too deeply characterized by the priorities of instrumental, results-based use (p. 6). Her theorizing is an intentional departure based in a recognition "that it is time to step back and reconceptualize the terrain of evaluation's influence" so as to arrive at a more "inclusive understanding" (p. 5). Henry and Mark (2003) concurred with Kirkhart's interpretation of both the conceptual and linguistic limitations of *evaluation use*: "we join Kirkhart in arguing 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 fragmented and marred the clarity of the field's theorizing. Instrumental use, symbolic use, process use, and conceptual use are all terms meant to connote some type of evaluation effect or utilization and are thus given the label "use."[80] However, their descriptors—process, symbolic, instrumental—do not fit together within one typology but actually describe disparate things:

1) The *source* of an evaluation's effect (like its evaluation process as in process use).
2) The *motivation* behind an evaluation's use (like the desire to legitimize a decision as a reason for evaluation in symbolic use).
3) The *form* of an evaluation's use (as with instrumental use) or the *form* of its effect (as with conceptual use).

[80] Kirkhart explains her interpretation of Weiss's (1981) suggestion of the term "use" over "utilization" to expand the field's articulation of evaluation's effects: "I strongly share her concern for selecting an accurate term that does not inappropriately constrict our understanding of the impact of evaluation; however, I disagree that the term use is a significant improvement over utilization. Not only are both terms instrumental and episodic, but they also imply purposeful, unidirectional influence…. One must step back from a narrow construal of use and rejoin earlier broad-based conversations. A broader construct than use alone is needed to represent this integration—one that does not privilege results-based use over influence stemming from the evaluation enterprise itself, one that does not chronologically limit our vision of the effects of evaluation, one that looks beyond the sightline of our intentions" (p. 7).

These descriptors constitute types of evaluation use that are not mutually exclusive from one another. Rather, they reflect discordant approaches (or scaffolds) for describing evaluation use that if harmonized perhaps could contribute to a fuller understanding of evaluation's effects.[81] The three dimensions of influence proposed by Kirkhart (2000) do not explicitly outline the varying *forms* (e.g., instrumental) that evaluation influence takes, nor do they describe different *motivations* (e.g., desiring legitimacy) that lead to evaluation influence. Her theory does not address some of the attributes emphasized in the literature on evaluation use. Nonetheless, Henry and Mark (2003) address both motivation and form in their theorizing on influence: their articulation of social betterment proposes an *ideal* motivation for the social change processes that evaluations inspire and that lead to influence. They also outline varying levels and types of evaluation influence—the former are described above, but the latter correspond to several different *forms* of evaluation influence (e.g., *priming*, *agenda setting*, and *diffusion*) within the varying levels (i.e., individual, intrapersonal, and collective).

In this study, employing the terminology of influence over use is largely about heeding a conceptual need. In light of the Western bias and limitations of evaluation theory as well as the unconventionality of ASER (Goodnight, Paper 2), I believe it is necessary to approach this analysis with a greater degree of conceptual fluidity in examining evaluation, its meaning, and its implications instead of searching for exacting terms expedient for definitively categorizing ASER's effects. I believe the broadness and integration that Kirkhart's (2000) theory of evaluation influence offers is important as is her reflectivity about the difficulty of negotiating time in explicating evaluation effects—she reveals that delineating use and influence could further be tied to whether the impact of an evaluation is seen as a "point-in-time event" or as an

[81] For example, Alkin and Taut (2003) point out that process use is not necessarily distinct from instrumental use because the process of an evaluation could be the source of an evaluation's instrumental use.

"open-ended process" of significance (p. 15). Quoting Cronbach's work, she suggests that evaluations' meanings are progressive "as 'part of the continuing accumulation of social knowledge'" (Chronbach, 1982, p. 318, as cited in Kirkhart, 2000, p. 15). I find the notion that an evaluation's influence is potentially an open-ended process contributing to the accumulation of social knowledge insightful in my analysis of ASER. Similarly, Henry and Mark (2003) appear to be arguing that evaluation influence is the confluence of various social change processes as links in a "causal chain" that ideally lead to social betterment (p. 305).

The Influence of Civil Society's Monitoring and Evaluation Efforts

A study of monitoring and evaluation (M&E) efforts facilitated by the Ghanaian NGO, the Social Enterprise Development (SEND-Ghana) Foundation, employed Henry and Mark's (2003) evaluation influence framework to explore district-level impact (Gildemyn, 2014).[82] During 2009-2010, SEND-Ghana monitored Ghana's National Health Insurance Scheme (NHIS) via "a decentralized network structure of District Citizen Monitoring Committees (DCMCs)" that had representatives from "various community-based organizations" and through "focal nongovernmental organizations (FNGOs), which 'host' these DCMCs" (Gildemyn, 2014, p. 511). SEND-Ghana's overarching purpose was "oversight and compliance" through taking a "nonconfrontational approach" to investigating the NHIS's "implementation and compliance with guidelines and objectives established by the government" and with "indicators…directly derived from the government program" (Gildemyn, 2014, p. 511). After validating their data with stakeholders, SEND-Ghana presented "aggregated findings at the national level in the form of an M&E report" that was launched with national media and followed by "policy dialogue

[82] Gildemyn (2004) perceives monitoring and evaluation as "two distinct but complementary processes" with "twin goals" of "accountability and feedback/learning" (p. 522). Also, see Mark and Henry (2004) for further explanation of the influence framework used in Gildemyn's study.

meetings at the district, regional, and national level" organized in conjunction with the FNGOs (Gildemyn, 2014, p. 511). The objective of these meetings was to facilitate direct conversation between the government and citizens. The ultimate aim of the study was to better understand the influence of these meetings on NHIS management and health service providers by identifying the "influence mechanisms" (i.e., change processes) occurring at the district-level as a result of SEND-Ghana's efforts.

In the study, SEND-Ghana's efforts—labeled as "participatory monitoring and evaluation" (PM&E)—are situated within a broader trend of civil society organizations (CSOs) independently monitoring government programs and policies in countries in the Global South (Gildemyn, 2014, p. 522). [83] Gildemyn (2014) identifies this trend as a phenomenon that has grown substantially over the past decade with "CSO-led M&E" efforts "burgeoning across the globe due to their increasing popularity with donors" and with the enlarged "availability of funds" from multilateral institutions (e.g., the World Bank) for such "social accountability initiatives" (p. 508). As part of a worldwide "aid reform agenda," CSOs are increasingly encouraged and funded to engage in M&E activities because it is assumed that CSOs are "better positioned" to "increase domestic accountability" at "the decentralized level because of their proximity to the grassroots" (p. 507). According to Gildemyn (2014), local CSOs in aid-receiving, postcolonial nations are seen as having the sociocultural competence required to both navigate data collection for M&E and foster public interest in weighing M&E findings and in pushing for greater accountability.

Interestingly, the first cycle of ASER was conducted in the last months of 2004 at the beginning of this period identified by Gildemyn (2014). ASER's architects highlight only

[83] Gildemyn (2014) clarifies that CSOs include "a variety of organizations" like NGOs, think thanks, and women's and farmer's organizations (p. 522).

national and local factors in their reasons for ASER's creation and in their explanations of how they devised its evaluation design (Goodnight, Paper 1). It seems SEND-Ghana's initial PM&E activities that tracked "pro-poor government programs and policies" began at roughly the same time (Gildemyn, 2014, p. 508). Based on the information provided by Gildemyn (2014) and the findings of my study, there is no indication that ASER or SEND-Ghana's M&E activities informed one another. Nonetheless, several values, strategies, and research methods of CSO-led M&E efforts appear closely aligned with aspects of ASER's design, process, and philosophy. For instance, the values underlying CSO-led M&E revolve around the concept of "answerability" to stakeholders (or the public at-large) in the form of "transparency," "justification" of policy action, and "enforceability"—the latter defined as creating "incentives for power holders to respond and/or take corrective action" (Gildemyn, 2014, p. 509).[84] In addition, the strategies and methods of CSO-led M&E include relying "on a range of monitoring tools to assess the implementation of a particular government program or policy," using "the gathered evidence in combination with a targeted advocacy strategy to hold government officials…accountable," and resorting "to 'soft' sanctions, such as media exposure, 'naming and shaming,' or dialogue…to pressure public officials…to take action" (Newell & Bellour, 2002, as cited in Gildemyn, 2014, pp. 508-509). The following sections on ASER will further shed light on the parallels between it and these CSO-led M&E values, strategies, and methods.

There is a strong case for the importance of studying new types of "social accountability initiatives" like CSO-led M&E in context of the evaluation field's theories of influence. In particular, as such M&E efforts continue to evolve in their objectives and worldwide

[84] The concept of being answerable is fundamental to social accountability (Christie & Alkin, 2013); for a detailed discussion of answerability in ASER, see Goodnight, Paper 2.

prominence, their implications both for stakeholders and for the evaluation field are largely not understood. Gildemyn (2014) convincingly makes this argument:

> Many of these earlier independent M&E initiatives have been ‘‘rebranded’’ as social accountability initiatives. There is a conceptual fuzziness surrounding social accountability initiatives, but in general, they can be defined as ‘‘citizen-led action for demanding accountability from providers’’ and/or public officials (Joshi & Houtzager, 2012, p. 146). Within the spectrum of social accountability initiatives, some focus on M&E through the use of public expenditure tracking, citizen report cards, and community score cards; some concentrate on transparency, policy advocacy, budget literacy, civic education, lobbying, or coalition building; and some combine several of these elements (McNeil & Malena, 2010, p. 8)….However, little is known about the outcomes of such initiatives in terms of their effectiveness and impact (Joshi & Houtzager, 2012; McGee & Gaventa, 2011). This paradox can be partially explained by the conceptual fuzziness, the scarcity of large-scale, comparative research on the topic, and a lack of understanding of how these initiatives work. (p. 508)

Perhaps, there is “conceptual fuzziness surrounding social accountability initiatives” not only because they are understudied, but also because they are a rather loose category of monitoring and evaluation efforts that have developed partially in response to global incentives (i.e., aid reforms, donor priorities, and increased funding) and partially, at least in some instances, organically in diverse places—even while being motivated by similar values and utilizing similar strategies and methods. Furthermore, it is possible that even the narrower category of “CSO-led M&E” (as compared with the broader category of social accountability initiatives) contains different monitoring and evaluation efforts. The differences between various CSO-led M&E efforts may reflect unique contextual and philosophical elements that are important in that they shape their respective designs, processes, and influences. This possibility only adds to the significance of studying how these various initiatives (i.e., M&E efforts) work. As the next section describes, this paper explores how the design and process of ASER create opportunities for various kinds of influence and contribute to the field’s theorizing of evaluation influence.

Study's Questions and Methods

This paper is based on a 10-month ethnographic case study of the process underlying ASER in India; my data draws heavily on field research (during 2014-2015) in villages, schools, and partner organizations' training venues with ASER Centre staff members and ASER participants who were involved in activities related to conducting ASER 2014 (see Goodnight, Paper 2). I spent considerable time in the ASER Centre central office in Delhi and in their state offices observing the usual activities of staff and the phases of the ASER evaluation process, principally in three states: Rajasthan, Tamil Nadu, and Manipur. My study relied on the following data collection methods: 1) semi-structured interviews, 2) observations, and 3) document analysis. My wider study (i.e., my ASER research beyond the scope of this paper's questions) examines ASER's purpose and goals, design and measurement choices, implementation, dissemination of findings, and influence by drawing upon the experiences and knowledge of ASER Centre staff. I also engaged ASER collaborators, community members, educators, government officials, and other members of the "policy-shaping community" (operating in India and beyond it) to provide valuable information on the local and policy context of the evaluation and perceptions of ASER's influence at various levels. The study resulted in roughly 90 interviews, over a thousand hours of observations (of ASER training, field data collection, monitoring, and dissemination of findings), and the collection of hundreds of documents (e.g., reports, data collection forms, training manuals, staff blog posts, and media articles).

I used a *vertical case study* approach, which is a multi-level analysis investigating an organization's influence upward through policy and bureaucratic structures (i.e., at state, national and/or global levels) and downward via relationships with local institutions (e.g., schools, village

councils), communities, and individuals. The purpose of vertical case study in this instance is toward understanding the evaluation's influence on individuals' participation and perceptions, on intrapersonal norms and communication, and on collective agenda-setting and action. Evaluation case studies have been considered effective in encouraging "evaluation use" and generating "a rich bounty of novel data" to support the discussion of evaluation's potential (Stake 1975 as cited in Shaddish, Cook, & Leviton, 1991, p. 271) and thus, a case study approach is evidently appropriate for exploring ASER's evaluation influence.

One of the big questions motivating the larger, ethnographic study of ASER is, *What influence is ASER having on educational change?* Here educational change may be interpreted as bringing about beneficial changes to the social conditions within India's education system and potentially beyond India—perhaps, this encompasses changes to educational beliefs, agendas, policies, planning, or practices at different levels. Questions that guide my inquiry into ASER in this paper include the following: How does the literature on evaluation influence guide analysis of ASER's effects? How does looking at evidence of ASER's impact challenge current theorizing on evaluation influence?

This paper, rather than providing a full accounting of ASER's influence in India and beyond it, utilizes the ASER evaluation to empirically explore Kirkhart's (2000) integrated theory of evaluation influence (and to some extent the levels and types of evaluation influence proposed by Henry and Mark). One aim of my larger study is to contribute to the field's understanding of evaluation influence, especially outside of Western contexts, via an investigation of how ASER aligns to the field's current theorizing about 1) the sources of evaluation influence, 2) the reasons for evaluation influence, and 3) the ways influence differs

from evaluation use. Eventually, this paper addresses the above-mentioned questions by answering another question: What is *design influence*?

Limitations

This paper has two limitations: first, it does not engage the several critiques of evaluation influence presented by Alkin and Taut (2003) or Nunnelly, et al. (2015). I offer a brief explanation of the issues with theorizing on evaluation use that prompted the development of evaluation influence, and I provide my own justification for choosing Kirkhart's (2000) influence framework for my analysis, but I do not relitigate the points made by these other texts. Given my paper's primary purpose of investigating how ASER informs theorizing about evaluation influence, such an endeavor would require a separate paper given the breadth of the discussion necessary to adequately address their critiques. Nonetheless, where directly relevant to my analysis of ASER, I draw upon and cite the work of Alkin and Taut (2003). Though certainly less commonplace than studies on evaluation use, scholars continue to employ and develop theories of evaluation influence because such frameworks provide for unique investigations of evaluation's meaning and effects (e.g., Morabito, 2002; Gildemyn, 2015). Secondly, this paper does not describe the history, methodology, structure, or process of the ASER evaluation to the extent that I have done elsewhere (see Goodnight, Paper 1; Goodnight Paper 2). Due to space constraints and the immensity of these aspects of ASER, it is impossible to be both thorough and stay within the parameters of a single paper. What I have strived to do in this paper is to provide enough detail and context about these aspects of ASER to create a substantial foundation for my discussion and theorizing in service of the above-mentioned research questions.

The Design, Process, and Philosophy of ASER

ASER's citizen-engaged evaluation model—including its goals, methodological design, and process—developed as a local response to educational challenges being witnessed in urban and rural areas of India (Goodnight, Paper 1; Goodnight, Paper 2). Pratham, India's largest educational non-governmental organization, created and piloted ASER in 2004 following its discovery that too many children attending its educational programs were not learning at expected levels despite their enrollment in school. Working in cooperation with government schools and local officials to strengthen the basic education of Indian children in the early 2000s, Pratham workers noticed that children struggled to read simple texts and complete basic math computations even after attending primary school for multiple years. By 2003-2004, Pratham began testing children, compiling test data, and sharing it with rural communities in "village report cards." Pratham discovered that disseminating test data in villages was powerful. Proof and clear articulation of children's learning problems proved vital to Pratham gaining local support for its interventions. While repeated findings from village report cards that children were not learning basic reading and math was distressing to Pratham workers, it was unclear how prevalent the issue was beyond the communities in which Pratham worked. In short, were children struggling to learn nationwide?

As Pratham was engaged in on-the-ground programming, the NGO's leadership was also participating in national policymaking efforts (see Goodnight, Paper 1). Consequently, Pratham's leaders cultivated a distinct outlook on the educational 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 (Goodnight, Paper 2). Pratham saw a need for annual, nationwide learning data that could be disaggregated to the state- and district-levels and that would provide systematic evidence of

whether children going to school in rural areas equaled children learning foundational skills. Eventually, Drs. Madhav Chavan, Rukmini Banerji, and Wilima Wadhwa created ASER's trailblazing methodology in response to these identified needs (Goodnight, Paper 1). The leadership of Pratham undertook ASER independent of the Government of India. ASER's three architects believed ASER could serve as a "proof of concept" to government officials that educational data could be collected reliably and inexpensively to inform policymaking, budgetary decisions, and school initiatives (Goodnight, Paper 1).

The main goals of ASER have remained the same since throughout its first decade, although some developments have been important to the evaluation's execution and indicators. In 2005, ASER volunteers first endeavored to find out if children were learning basic skills in reading and math, and Pratham's leadership sought to evaluate those children's learning levels in context of information that volunteers collected about children's enrollment and households. In addition, ASER volunteers collected data about the basic condition and functioning of government schools. After a few cycles of ASER, Pratham was awarded funding from Google in 2008 to establish an independent NGO, called ASER Centre, which would assume all responsibilities for conducting ASER while Pratham would remain an affiliated organization that focused principally on educational advocacy and providing learning programs (i.e., interventions). After the passage of India's Right to Education Act (RTE) in 2009 (Government of India, 2009), the school surveys administered during ASER included specific indicators to evaluate RTE's implementation. The Act provided ASER Centre with further indication as to what measures the GOI thought were credible indicators of school quality. However, RTE's language did not specify how children's learning should be measured, neither validating nor

invalidating ASER's approach to assessing children's learning with basic metrics for reading, arithmetic, and English language proficiency.

Main Features of ASER's Design

Appendix B outlines the many features that constitute ASER's evaluation design, including its use of simple learning assessments, its household-based approach to testing, and its engagement of volunteers as field data collectors and partner organizations as trainers and facilitators of the evaluation (see Goodnight, Paper 1; Goodnight, Paper 2). The ASER architects' choice of one-page reading, arithmetic, and English language tests that each assess four levels of competency up to a grade-level 2 or 3 proficiency constitute a significant feature of ASER's design. The simplicity of the tests enhances the capacity of thousands of ASER volunteers to assess children's learning levels reliably. Moreover, tests are designed to straightforwardly assess foundational literacy and numeracy skills in a manner so even illiterate parents can decipher the tests' levels and meaning. This fosters ASER's ability to raise awareness about children's learning within rural communities where many adults have limited experience of schooling. As mentioned earlier, household-based versus school-based testing enables ASER's representation of children who are out of school and enrolled in various types of schools in its data (Goodnight & Bobde, unpublished). It also facilitates collection of data on home factors that have been shown in education literature to influence schooling (e.g., household assets and mother's education).

The citizen-engaged evaluation model is largely defined by its involvement of volunteers and partner organizations from across India. Volunteers—typically young pre-service teachers, university students, or NGO workers—are recruited from each district to conduct data collection for ASER in villages. After being recruited and trained, the volunteers travel to sample villages

for village mapping, household surveying and testing, and school surveying. Using local volunteers to conduct the survey helps ASER to transcend vast cultural and linguistic differences across rural India and increases the trustworthiness of the data collected. Moreover, using volunteers minimizes costs and engages young people in educational issues and exposes them to social research methods (Goodnight, Paper 2). Partner organizations—usually government teacher training institutes (called District Institutes of Educational Training or DIETs), universities, or non-governmental organizations—help facilitate ASER at the district-level. The involvement of a partner institution for each district in every state helps minimize the costs and logistical hurdles of conducting ASER for ASER Centre. Partner organizations provide personnel, space, and other forms of assistance that the Centre cannot itself maintain in the 500+ districts that it conducts ASER. Moreover, partner organizations have vital local knowledge (regarding communities, language, culture, security issues, and so on) that makes it possible to execute ASER in districts with difficult terrains, social tensions, and other challenges.

The ASER Process

The execution of ASER encompasses a massive range of activities—training sessions, travel to villages, surveying and testing, monitoring, data recheck and aggregation, data analysis, and report writing and printing. Its evaluation process involves the coordination of not just volunteers and partners, but also hundreds of ASER Centre's own staff. The phases of the ASER process are depicted in Figure 4. ASER's annual activities begin in June with forming partnerships and extend into March with the dissemination of ASER findings to various stakeholders at the national, state, and district levels (see ASER Centre, 2015 and Goodnight, Paper 2 for more in-depth explanations of ASER's process). Phases I-IV necessitate the heaviest participation of ASER volunteers, partners, master trainers as well as ASER state and central

staff members. The roughly 100-day timetable from "field to report" (i.e., late September to early January) is an important feature of ASER's design in that its up-to-date findings can inform decision-making in nearly real time.[85]

To understand the phases of ASER's evaluation it is crucial to comprehend the tiered structure of its training, data collection, and monitoring process (see Figure 14). ASER is directed from the top by the Delhi-based ASER Centre national team members, called Research Associates, who oversee state teams assigned to them and help orchestrate logistics between states and the national office. Meanwhile, state team members are located year-round in ASER state offices across where they liaison with partner organizations and they supervise everyone who executes the ASER evaluation *within* their respective states, including master trainers (MTs) and volunteers. Across India, roughly 1,000 master trainers are hired each year to facilitate ASER. MTs coordinate logistics for the ASER evaluation at the district-level, including leading training workshops with help from partner organizations. MTs also monitor the volunteers during field data collection and conduct data quality rechecks for their districts. Meanwhile, volunteers conduct the mapping, surveying, and testing in sample villages, households, and government schools.

[85] Because of India's scale and complexity, the fact that ASER data are current (i.e., reported within a few months of its collection) and annual represent a substantial departure from the large-scale data collected and reported by the government entities. For example, see ASER Centre (2015, pp. 316-319) for a comparison of ASER and the government's National Achievement Survey (NAS).

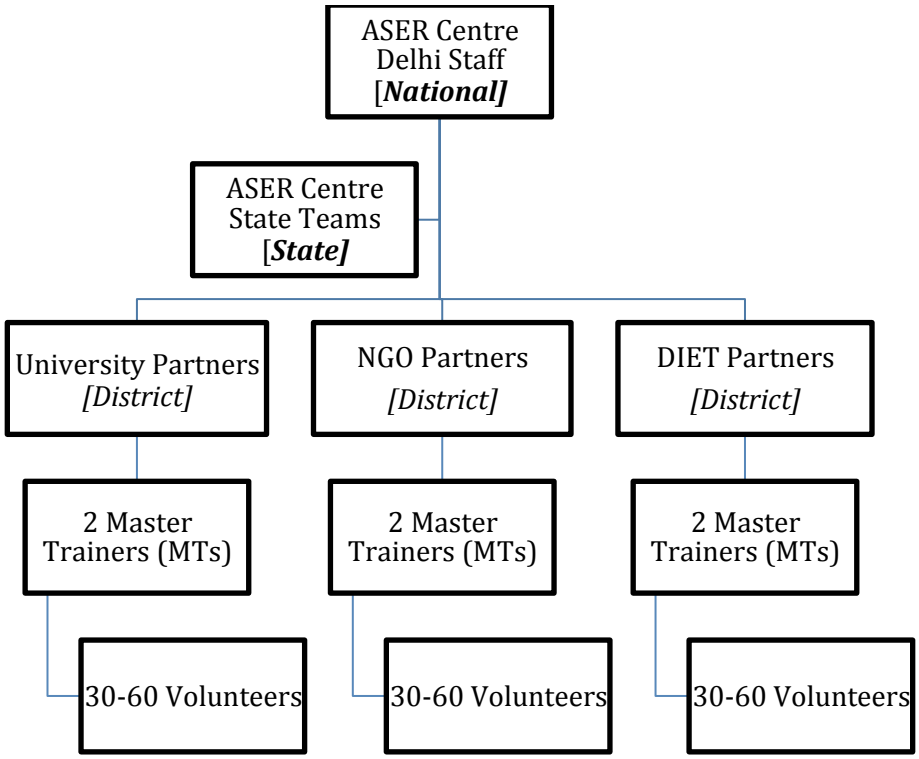

Figure 14. The tiered structure of the ASER process

ASER's Philosophy, Values, and Social Goals

The citizen-engaged aspect of ASER is born not just out of a logistical need to overcome the data collection issues of scale, cultural diversity, security, and cost, but also out of a philosophical stance on the need for popular engagement in education (Goodnight, Paper 1). The Centre's vision for ASER is a democratic one defined by "shared responsibility" (Chouinard, 2013, p. 239) and "social accountability" (Christie & Alkin, 2013, p. 14).[86] ASER directly advocates public ownership of social systems that affect the everyday lives of ordinary people (Goodnight, Paper 2). On a large scale, ASER is striving to realize evaluation's democratic potential by encouraging knowledge sharing and deliberation to advance the public good (House

[86] Christie and Alkin (2013) define social accountability as "true accountability [which] requires 'answerability'—that is, those responsible must be held accountable" for the quality and equity of social programs (p. 14). See Goodnight (Paper 2) for a more detailed discussion of how ASER exhibits social accountability and advocates shared responsibility.

& Howe, 2000). ASER's evaluation is purposefully designed to simultaneously motivate policy change at the top and stimulate individual and community action at the grassroots (Goodnight, Paper 1; Goodnight, Paper 2). Both long-term goals (i.e., policy change and grassroots action) are intrinsic to ASER's "theory of change," and these goals rely on the interim objectives of generating public dialogue at the national and state levels and awareness at the district and community levels (Results for Development Institute, 2015, p. 12).

ASER is a form of social systems evaluation that creatively integrates elements of monitoring, policy evaluation, and program evaluation to achieve its long-term goals of policy change and grassroots action. ASER addresses priorities outlined in India's Right to Education Act and incorporates indicators outlined in RTE to prompt widespread debate with its findings on the status of government schools and children's learning under the policy (i.e., monitoring) as well as to raise vital questions about the logic and adequacy of RTE prescriptions (i.e., policy evaluation) (Goodnight, Paper 1). The Centre leadership hopes quantitative evidence illuminating RTE's degree of implementation and its overall success thus far can create the public pressure necessary to motivate beneficial planning and policy changes (Goodnight, Paper 1). As a consequence of its policy goals, ASER's design integrates elements of both monitoring and policy evaluation.[87]

[87] It can be difficult to definitively separate "auditing, monitoring and inspection" efforts from those of policy evaluation as the former "are all capable of generating data that can be used in [the] evaluation" of a policy, but ultimately, evaluation "goes well beyond these activities, by seeking not just to describe how a programme [i.e., the policy] is operating but also aiming to explain the underlying logic" of it (Clarke & Dawson, 1999, p. 7). ASER undoubtedly fulfills several criteria conventionally ascribed to monitoring in that it is "an on-going activity" (i.e., annual) that relies upon the "identification of measurable indicators" that produce data that are, in part, used to "assist in budget decision-making" (Clarke & Dawson, 1999, p. 7). ASER also utilizes "outcome-related performance indicators" to evaluate a public sector system "in the interests of public accountability" (Clarke & Dawson, 1999, p. 7). However, ASER constitutes evaluation in that its goals and activities go beyond mere performance measurement because ASER is striving to raise "serious questions of public direction" in relation to the government's provision of education and to stimulate critical, widespread debate about whether what RTE is doing is "a good thing" or whether its priorities are adequate (Schwandt and Gates, 2016, p. 67)—these activities that examine and challenge the logic of RTE (and even question smaller education policy decisions made by the government, such as cutting grants for teaching and learning materials) all constitute endeavors more aligned to

At the same time, ASER aims to raise awareness at the grassroots in ways that reflect elements fundamental to participatory program evaluation, which seeks to engage a diversity of stakeholders in the evaluation process. Participatory program evaluations are "sensitive to local contexts" and claim values of "local ownership, empowerment, use of findings, organizational and individual learning, and program improvement" (Chouinard, 2013, pp. 237-238). Design choices made for ASER reflect sensitivity to local contexts and promote individual learning and local ownership of government schools (see Goodnight, Paper 1; Goodnight, Paper 2). In particular, the decision to use local volunteers as ASER field "surveyors" (i.e., data collectors) reflects the above-mentioned considerations and values in that the choice was governed by a desire to competently navigate India's linguistic and cultural diversity (i.e., contexts) while promoting widespread awareness about learning as an important educational outcome. Dr. Wadhwa explains:

> We realized that it [i.e., ASER] *had* to be done in regional languages. You can't take people from Delhi, who will go and test in Telegulu in Andhra Pradesh,[88] so it had to be some local [people]. (W. Wadhwa, personal interview, May 15, 2015)
>
> At some level if you are going to build *awareness*, if the goal is to shift the focus from inputs to outcomes…you have to build that awareness at *all levels*. It is not just at *central* policymaking. It is *state* policymaking. It is parents. It is the *common* man, right? That is where the *citizen-led* kind of nature came out of it. (W. Wadhwa, personal interview, May 15, 2015)

The "citizen-led" nature of ASER emerged from an objective of engaging people at all levels of society in the cause of primary education and specifically, the architects of ASER valued how

---

policy evaluation. Also, uses of ASER data to pose larger "cause-and-effect questions" (Chelimsky, 1985 as cited in Clarke & Dawson, 1999, p. 5) and to put forward hypotheses about the relationships between government and private school provision, household factors, and children's learning surpass commonplace definitions of monitoring (For an example, see ASER Centre, 2015, pp. 19-21).

[88] Telegulu is a language typically spoken in Andhra Pradesh and Telangana.

ordinary citizens could benefit in their learning and understanding by being directly involved in data collection:

> Fifty percent of children in India, like ASER says, cannot read a Standard II-level text. Yes, these are shocking numbers. But, what is *far* more shocking is to go into that village and to sit down with that child; you have recorded that the child goes to school; the child is in grade 5, and you see the child *struggling*. Now, *nothing* brings the problem home until you *actually see* that. It's like, you know, it is one thing to talk about malnutrition, but it is another thing to see that child with that *potbelly*.... When you actually see that, there is nothing that brings the problem home [as that does]. (W. Wadhwa, personal interview, May 15, 2015)

Through her malnutrition analogy, Dr. Wadhwa conveys that the ASER architects understood a dual reality—one of quantitative data and one of firsthand experience. As fieldworkers and researchers, they recognized that numbers were an inevitable abstraction of children's learning struggles, but firsthand experience humanized these issues, this reality that numbers attempted to unveil. In short, they believed if they could involve ordinary citizens in fieldwork as "volunteers," ASER could build an unparalleled grassroots awareness across India about the status of its primary education system. However, ASER's end objective was and is not awareness, but action. Through firsthand engagement in testing children and talking to families in villages about schooling, the ASER architects wanted ASER volunteers and villagers to evolve in their sense of investment in children's education and ownership of the Indian public school system. To date, ASER asks young people to donate their time toward traveling to villages in their district and investigating what children can do in reading and math, believing that such an experience can essentially alter citizens' commitment to improving primary education. Therefore, informed by its values of citizen participation, deliberation, social accountability, and shared responsibility, it is ASER's novel synthesis of elements of monitoring, policy evaluation, and program evaluation in its design for evaluating a large-scale social system (i.e., primary education) that defines its citizen-engaged model.

Discussion

The Sources of ASER Influence

Recalling Kirkhart's (2000) integrated theory of influence, there are two kinds of evaluation influence based on the dimension of "source"—*results-based influence* and *process-based influence*—that inform this analysis of ASER. Results-based influence relates to the effects of ASER findings (i.e., data, conclusions, and recommendations), which could be on policies, initiatives, practices, public dialogue, and so on. In contrast, process-based influence relates to the effects of the ASER process on institutions and individuals who participate in ASER. Consequently, ASER may have process-based influence on partner organization personnel, master trainers, volunteers, school headmasters and teachers, parents, community members, and ASER central staff and state team members. According to ASER's theory of change, which informs its evaluation design, the two goals (i.e., policy change and grassroots action) are intended to emanate from two different sources of evaluation influence.

Results-based influence in ASER's theory of change. In ASER's theory of change, the first goal of policy change is conceived as resulting from ASER findings. In other words, policy change is supposed to reflect ASER's results-based influence. Accordingly, the degree to which ASER instigates policy changes in education at the upper levels of government is chiefly related to the degree of influence of ASER findings. The means by which ASER findings are believed to affect policy change are via prompting public deliberation and public pressure on government officials. Thus, ASER's potential influence on policy can be intermediately evaluated through asking how and to what extent ASER's results prompt public dialogue and pressure on government toward making policy changes. As Gildemyn (2014) highlights in CSO-led M&E, public dialogue and pressure are often generated through media coverage of findings.

Consequently, an interim objective (or "end-of-cycle" objective) in promoting ASER's results-based influence on policy is prompting public dialogue and pressure via media coverage of its findings. The long-term goal is government making policy changes in education, including planning and budgetary reforms. The ultimate goal (or greater social purpose) of ASER's findings can be conceived as social betterment through an improved education system as a result of constructive policy change. Figure 15 illustrates the intended pathway of ASER's results-based influence according to ASER's theory of change.

Figure 15. The intended pathway of ASER's results-based influence on policy (via media).

An alternative pathway for ASER's results-based influence is reflected in Figure 16. In the influence of ASER findings, it illustrates different interim objectives of government officials' enhanced knowledge and decision-making but the same long-term goal (i.e., government making policy changes in education) and the same greater social purpose of an improved education system. Instead of the pathway of ASER's results-based influence beginning with the coverage of its findings in the media, this pathway begins at the end of ASER's cycle with the dissemination of its findings to key officials by mailing them or giving them the report (i.e., directly sharing with them ASER's published data and recommendations). This alternative pathway for results-based influence reflects a somewhat different assumption about the motives or interests of officials—that they will be amenable to consuming the ASER findings, to

assimilating them into their existent knowledge of the education system's issues or dynamics, and subsequently, to considering these findings in their decision-making and eventually, in their attempts to change educational policies toward an improved system. In the scenario depicted by Figure 16, government officials do not need public pressure to enact change, rather they can be convinced to change policies based on their enhanced knowledge of the educational outcomes or problems under current policy.

Figure 16. The intended pathway of ASER results-based influence on policy (via decision-making).

Process-based influence in ASER's theory of change. The second goal in ASER's theory of change, which is grassroots action, is contrastingly theorized as stemming from ASER's process. The ASER process is intended to generate individual learning and awareness about education and group, organizational, or community awareness. This means community-level action undertaken by either groups or individuals reflects ASER's process-based influence. The degree to which ASER stimulates grassroots action is equated with the degree of influence of ASER's process on those directly involved in it. ASER's grassroots influence can be intermediately evaluated through asking, for example, how and to what extent ASER's process prompted individual volunteers' awareness, partner organization awareness, or a group of villagers' awareness about the government primary school system or children's learning toward spurring further deliberation and action. An interim objective (maybe "end-of-cycle", but could

be “long-term” according to Kirkhart’s time sub-divisions) is prompting local deliberation, organizing, and/or intermediate actions around education. The long-term goal (or longer-term) is ongoing, well-orchestrated action and a prolonged, future commitment to the cause of education. The ultimate goal (or greater social purpose) of ASER’s process can be viewed as social betterment through improved, sustained or institutionalized local educational practices as a result of constructive action. Figures 17 and 18 illustrate two intended pathways of ASER’s process-based influence—one related to a trajectory of individual action and the other group (i.e., coalitional) action—according to ASER’s theory of change. Figure 18 illustrates that group action may also be the result of an interim change of increased individual awareness that leads to enhanced group awareness.

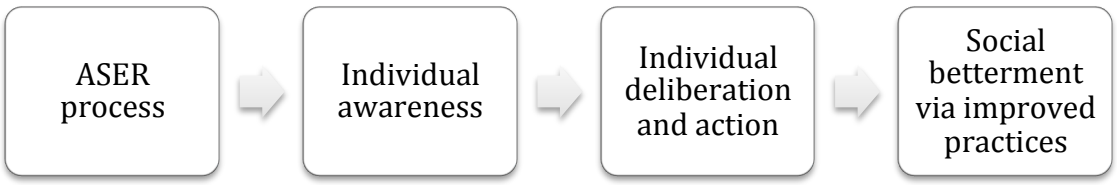

Figure 17. The intended pathway of ASER's process-based influence on individual grassroots action.

Figure 18. The intended pathway of ASER's process-based influence on coalitional grassroots action.

Model pathways versus actual influence. In summary, ASER’s theory of change, which shapes its evaluation design, includes two goals that emphasize different plans for achieving

results-based influence and process-based influence. The goals of policy change and grassroots action actually demonstrate the explicit intention of results-based influence and process-based influence in ASER's evaluation design. These intended influences are a significant theoretical aspect of ASER's design as they present ideal (i.e., model) pathways for achieving ASER's grander purpose—that is, for realizing ASER's specific vision of social betterment.[89]

These ideal pathways for results-based influence and process-based influence (depicted in Figures 15-18) are oversimplified and excessively linear in comparison to the probable messiness of how influence actually unfolds in ASER and also according to Kirkhart's (2000) characterization of influence that includes effects that are "multidirectional, incremental, unintentional, and noninstrumental, alongside those that are unidirectional, episodic, intended, and instrumental" (p. 7). In reality, it may also be difficult to always decipher which source of the evaluation (i.e., results or process) has prompted the effects or observed changes that lead to evaluation influence (Gildemyn, 2014). For instance, an ASER master trainer or state team member might be simultaneously inspired to take action within their community as a result of participating in ASER and knowing what the ASER evaluation has found in terms of children's learning levels (see the experiences of Fazal in Goodnight, Paper 2). Also, it is plausible that disentangling the observed, episodic uses of ASER findings or processes (typically labeled as "instrumental use") from the comparatively fluid, comprehensive idea of results-based influence or process-based influence is often not possible or desirable.

The potential relationships between use and influence can be explored through some concrete examples. For instance, ASER data has been cited for several years in Government of

[89] ASER's vision of social betterment is well articulated in its vision statement (see Goodnight, Paper 2): "When ordinary people learn to measure what affects their lives, they can communicate with each other across villages, states, nations, and continents, to identify and understand their problems, take steps to resolve them, and change the world for the better*"* (ASER Centre, n.d.b).

India's Ministry of Finance's Economic Survey reports (see Government of India, 2014a).[90] Nevertheless, it is important to consider the meaning of this use of findings by studying the extent to which it reflects something more in accordance with ASER's theory of change—such as prior, ongoing, or future use in GOI's decision-making—rather than isolated or episodic use. In other words, does publicly citing ASER findings in the report reflect ASER's broader influence on knowledge and decision-making, and will it lead to long-term positive and substantive policy changes in the allocation of resources to the education system? These are important topics for further study. Moreover, one might wonder if the GOI Ministry of Finance's citation of ASER findings influences the behaviors or thinking within other government entities, like India's Ministry of Human Resource Development (MHRD). For example, there is an open question as to whether MHRD's stance on the credibility or usefulness of ASER findings may be influenced by the frequent use of ASER findings in the Economic Survey by the Ministry of Finance. The underlying question is regarding whether episodic, unidirectional use of ASER findings in one government report is an important change process that prompts similar uses elsewhere in the Indian government. If use of ASER findings stimulates incremental, multidirectional, and noninstrumental changes, it could signify something more lasting and potentially momentous to achieving social betterment through ASER, which is intrinsically tied to fostering a "measurement culture" within India (Goodnight, Paper 2). For comparison, it is helpful to examine an alternative example of instrumental use relating to the ASER process. One could inquire whether the ASER process translates to ongoing process-based influence for an ASER-partner NGO that opts to use ASER basic learning tests in its own intake process for children participating in one of its programs. In determining whether this use is instrumental,

[90] ASER findings have been discussed in the following years of the Economic Survey: 2009-2010, 2010-2011, and 2013-2014. Additionally, ASER results have been cited in the Government of India Planning Commission's approach papers for their 11th (2007-2012) and 12th (2012-2017) five-year plans.

episodic, and isolated or whether it signifies a longer-term influence, one could examine if ongoing use of the ASER tests has observably influenced how the NGO's programming is conceived, implemented, or designed.

In summary, to determine whether ASER's model pathways of influence translate to actual influence (in the way its theory of change has intended), it is important to conduct empirical study of the actual effects of ASER. Future study on the actual results-based and process-based influence of ASER is the next step in my investigation of the data that I collected from this broader research on ASER. In particular, through interviews, documents, and observations, I will be looking for evidence of results-based and process-based influence as articulated by ASER participants and/or observed during the implementation of ASER's process as well as in documents and artifacts after the ASER cycle has concluded. Actual evidence related to ASER's results and process can facilitate an exploration of the alignment between these model pathways and the influence resulting from how ASER actually unfolds (Goodnight, Paper 1).

Revisiting the dimension of time. One challenge that was a commonly repeated theme in my interviews with ASER Centre's directors was the challenge of knowing in what timeframe the organization should follow-up with volunteers, a partner organization, a community, or a government official to judge the extent of ASER's impact (whether emanating from its findings or its process). Influence can "evolve over time into extended impact" (Kirkhart, 2000, p. 16), but it is difficult to determine what is a reasonable amount of time to let pass before proactively trying to investigate impact—not to mention for ASER the difficulty of following up with 25,000 volunteers, over 500 partner organizations, and 16,000 villages annually. Despite this challenge of time and scale, ASER's long-term influence is a significant issue in judging ASER's overall

meaning. As Kirkhart (2000) states: "Although influence during the process of evaluating and reporting results is important, the most powerful impact of the work may not yet have emerged or be visible in that time frame, lying instead in a future context" (p. 16). It seems that because ASER is a particularly intense process for its participants, revisiting the possibility of impact at a future date is logical as participants may take several months or years to assimilate and employ all that they have learned. Such investigation of ASER is beneficial in contributing to the evaluation literature: "the significance of long-term influence is well recognized…[yet] Shulha and Cousins (1997) found it to be noteworthy in its absence from empirical studies of use" (Kirkhart, 2000, p. 17). In taking a long view of ASER's potential influence, I discovered through interviews, informal conversation, and document analysis that a different type of influence was emerging—an influence that up until this moment had not been identified in the literature on evaluation influence. In looking at ASER's increasing reputation globally, it became apparent its evaluation design was its main source of influence with a global policy-shaping community that was eager to find solutions for measuring learning at a large-scale and with civil society stakeholders located in other nations of the Global South who wanted to similarly foster awareness around learning issues in their countries and increase accountability regarding the provision of primary education. Regarding the aims of the global policy-shaping community, generating national data on children's learning is a key focus for monitoring progress on the United Nations' fourth Sustainable Development Goal: "ensure inclusive and quality education for all" (United Nations, 2016).

The Global Influence of ASER's Design

As the theme of its October 2016 annual conference in Atlanta, Georgia, the American Evaluation Association (AEA) 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." As explained above, ASER is a model for citizen-engaged evaluation of social systems; ASER's design represents a synthesis of program evaluation, policy evaluation and monitoring. ASER's central theory of achieving social betterment is via broad citizen participation in monitoring learning and schools paired with a responsive government that exhibits a commitment to social accountability, which is shared with the wider citizenry who also takes responsibility for the quality of the education system. The approach of ASER's model is hierarchical (i.e., tiered) and participatory—engaging local volunteers and partner organizations but directed from above by the Centre's state and national staff members. The methods are village mapping, village surveys, government school surveys, household-based surveys, and household-based tests measuring children's basic learning levels in language and math.

The integration of the above-mentioned features pertaining to ASER's theory, approach and methods constitute ASER's design, which emerged organically to evaluate educational issues specifically in the Indian context (Goodnight, Paper 1). ASER has since achieved the status of a transnational model with its key design features being adopted by similarly-minded non-governmental and development organizations in other nations: MIA in Mexico; Jàngandoo in Senegal; Bɛɛkunko in Mali; SCALE in Ghana; LEARNigeria in Nigeria; Djangirde in Cameroon; Uwezo in Uganda, Kenya, and Tanzania; TPC in Mozambique; IID/BRAC Survey in Bangladesh; and, ASER in Pakistan. The features unifying the M&E efforts across these countries are

- Testing "basic reading and math competencies"
- Conducting testing and surveying "in households," "not schools"
- Conducting testing "orally and one-on-one"

- Creating an evaluation that is "statistically representative"
- Executing an evaluation that is "independent" (i.e., "organized by civil society")
- Testing "basic competencies" with only a "small number of items" for all children "regardless of age or grade-level"
- Engaging "citizens" in the "data collection process" instead of just "sharing data after the assessment"
- Aiming data at a "broad audience, not just authorities and policymakers" (Results for Development Institute, 2015, p. 11).

Each one of these features is intrinsic to ASER's design and was an intentional choice made by its architects during its initial development (see Goodnight, Paper 1). Recently, the People's Action for Learning (PAL) Network formed to support cross-country efforts to employ the ASER citizen-engaged evaluation model (called "citizen-led assessments" by the PAL Network and others) with financial support from the Hewlett Foundation.[91] A 2015 report produced by Results for Development Institute found that ASER had "contributed to an increased focus on learning outcomes in global discourse and agenda-setting" by "providing evidence of the seriousness of the learning crisis" and "demonstrating how a low-resource model can be used to assess learning on a national scale" (p. 4). ASER overcame several methodological challenges in creating its model of evaluation in order to produce reliable, independent data on basic learning levels in the linguistically and culturally diverse contexts of rural India (Goodnight, Paper 1). In doing so, ASER has created a model for large-scale educational evaluation that has evidently resonated with people trying to generate national learning data in other national contexts and

[91] An interview with a program officer from the Hewlett Foundation revealed that the officer believed she had created the term "citizen-led assessment" to describe ASER and that it gained momentum. Elsewhere, I argue that identifying ASER as only an assessment narrows the conceptualization of its model and its implications (see Goodnight, Paper 2).

thus, ASER has now "inspired the development of several citizen-led assessments in other countries" (Results for Development Institute, 2015, p. 10).

The ASER evaluation model continues to garner interest from multilateral institutions, think tanks, and development bodies like UNESCO and the Brookings Institution, which together founded the "Learning Metrics Task Force" that was co-chaired by Rukmini Banerji of ASER Centre. The Task Force determined that in context of global goals around Education for All, "assessment was a public good" and more specifically, 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 the evaluation of national education systems presents a development solution for producing (independently, and at a low cost) the large-scale data seen as necessary for closing the 'global data gap on learning outcomes' and for providing information crucial for ameliorating the current 'crisis' of 'poor quality education,' which is believed to be 'jeopardizing the future of millions of children and youth across high-, medium- and low- income countries' (Learning Metrics Task Force, 2013, p. 9). In addition, academics affiliated with the above-mentioned institutions or based at universities outside of India increasingly study ASER and cite ASER data, sometimes expressing either concern or support for its model's proliferation (e.g., Alcott & Rose, 2015; Barrett, 2011: Pritchett, 2013; Winthrop & Simons, 2013)

The *Design Influence* of ASER

Through my research on ASER, I have identified an apparent limitation of Kirkhart's (2000) integrated theory of evaluation influence: the oversight of an evaluation's theoretical underpinnings or design as a potential source of influence. Kirkhart's theory appears underdeveloped in its consideration of an evaluation's design as an influential model for future

evaluations. Rather, her theory locates only two possible sources of an evaluation's influence: an evaluation's process and results. Investigating the relationship between evaluation design and influence can increase understanding of an evaluation's broader long-term implications and meaning. Absent from Kirkhart's (2000) three dimensions is an explicit accounting of the evaluation's methodological design, including its underlying philosophy and goals as factors in its influence. While an evaluation's design (emanating from its philosophy, goals, and methodology) certainly shapes the entire evaluation process, the design is not the same thing as the process. Moreover, the dimension of "intention" captures whether the effects of an evaluation align with its initial goals, but this dimension doesn't speak to how an evaluation's initial goals can determine—either facilitate or impede—its subsequent influence, which is a complex matter of theory (i.e. discerning the relationship between causes and effects), not merely intention (i.e., motivation). Kirkhart's (2000) model is pioneering in its creation of a framework for studying evaluation influence, and it has enabled me (through this research on ASER) to identify how an evaluation's design itself can be a source of substantial and lasting influence.

My research on ASER suggests that some evaluations are so pioneering in their design that their model of evaluation—that is, their evaluation *design* that unifies the evaluation's theories, approaches, and methods—becomes that evaluation's defining influence. Figure 19 displays how the evaluation's design includes the intended evaluation process and results,[92] which have been previously identified by Kirkhart as sources of evaluation influence. In sum, my definition of *design influence* is the influence of an evaluation's design (its theory, approach, methods, process, and type of data) as an innovative model for the development and conduct of further (future) evaluations. Although an evaluation's design includes its intended process and

[92] Regarding the evaluation's results, I mean the evaluation design encompasses the *types* of data that are intended to result from the evaluation's indicators or methods, so the evaluation design is inclusive of the evaluation's findings unless the types of data or kinds of findings that emerge are unintended by design.

results, it is not equal to the actual process and results of the evaluation. The design is something quite beyond just process and results as it includes a more holistic engagement with philosophy, theories of change, goals, and methodologies.

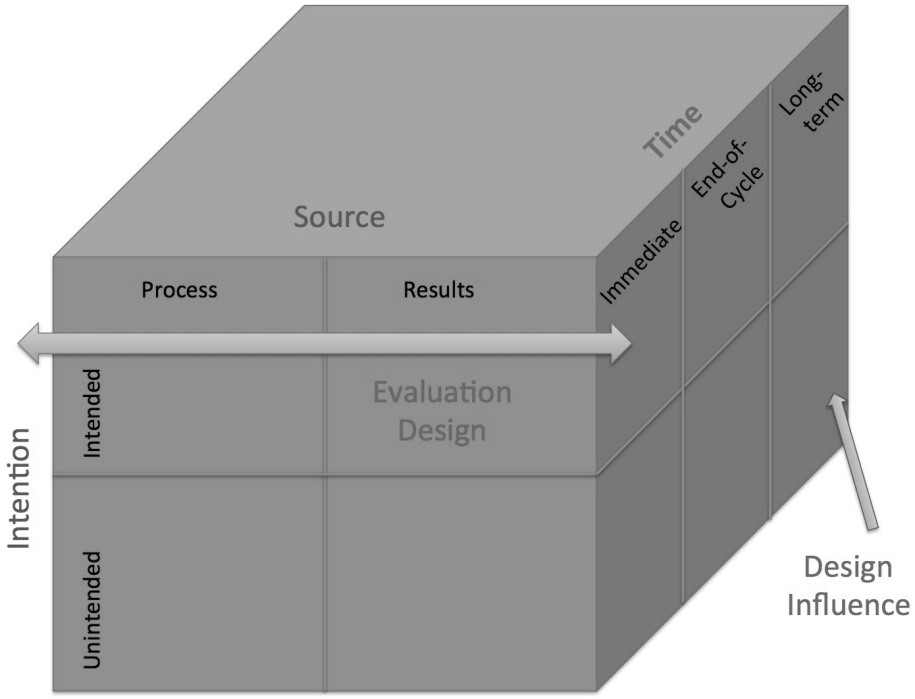

Figure 19. Revisiting Kirkhart's diagram of her integrated theory of influence. Adapted from "Reconceptualizing Evaluation Use: An Integrated Theory of Influence," by K. Kirkhart, 2000, *New Directions for Evaluation*, 88, p. 8. Copyright 2000 by Jossey-Bass Publishers. Adapted with permission.

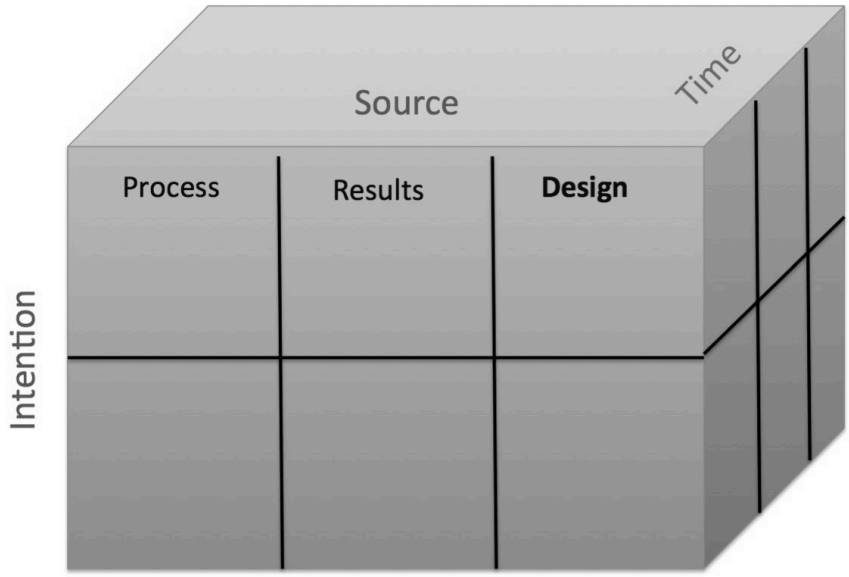

Figure 20. Revised model of evaluation influence with design influence. Adapted from "Reconceptualizing Evaluation Use: An Integrated Theory of Influence," by K. Kirkhart, 2000, *New Directions for Evaluation*, 88, p. 8. Copyright 2000 by Jossey-Bass Publishers. Adapted with permission.

Figure 20 depicts a revised model of Kirkhart's (2000) integrated theory of influence with design influence included as a source of evaluation influence. As the arrow in Figure 19 suggests, the design influence of ASER based on the analysis conducted for this paper appears to be incremental (i.e., long-term) and "unintended;" future analyses of my ASER data may identify other instances of ASER's design influence that are, for example, episodic or immediate. For now, I identify design influence as principally unintended and long-term: the architects of ASER discussed creating its evaluation design only in response to the needs that they identified within the Indian education context, so the influence of its design was not part of its initial purpose; the architects wanted to meet certain goals that they had for policy change and grassroots action in the rural contexts in India within which they had been working (Goodnight, Paper 1). AEA's definition for evaluation design importantly includes the phrase "to achieve a set of intended purposes in a specific context," which I interpret as an evaluation design's specification for investigating a particular evaluand in an actual context with a clear purpose. What is curious

about the influence of ASER is that once its evaluation design becomes a "model" or once its evaluation has *design influence,* its theories and methodological choices become unmoored from the context for which the ASER evaluation design was originally created—namely, rural India. The other curious thing about ASER's design influence is that it transcends the audience (or stakeholders) for which ASER was created, which was already quite broad to begin with because the original audience included all ordinary citizens of India. The design influence of ASER seemingly relates to the concept of *diffusion*, 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, 2003, p. 305). However, the ASER form of influence is design diffusion where the practice(s) being adopted are not what is under evaluation but is the model of evaluation itself. The design influence of ASER reflects design diffusion on a global policy-shaping community and on groups of stakeholders in other countries, which the ASER evaluation never intended to reach or have influence upon when it was created. This is ironic in some ways as ASER did begin as a "proof of concept" with an immense attention to the detail of its design, but for the purpose of persuading the Indian government of the design's efficacy, not stakeholders outside of India (Goodnight, Paper 1).

Theorizing from this research on 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, 2) how it overcomes challenges to conducting evaluation in a specific context (or certain type of contexts), 3) how its philosophical underpinnings reshape the purpose of evaluation, and 4) how its approach uniquely involves various stakeholders in the evaluation. The influence of ASER's design in new countries moves beyond any singular cycle of ASER's evaluation in its original context of India. ASER has design influence apparently

because it provides solutions for future evaluations in these different spaces. Perhaps, the ASER model will eventually be adapted for use beyond its original evaluand of basic learning and primary education. Nonetheless, translating evaluation designs between locales can be problematic as the merit of ASER's citizen-engaged evaluation model for rural India is two-fold: 1) it reflects an understanding of the challenges of collecting accurate data on learning and education in Indian villages and government schools and 2) its methods and process are shaped by a strong commitment to its theory of change, which grounds its meaning in something beyond mere monitoring, data collection, or assessment—ASER is grounded in a philosophy of social accountability and shared responsibility through its vision of social betterment. Further research on how the ASER model has been adapted in these other national contexts and what influence its citizen-engaged model is having on different education systems and stakeholders is important to undertake. My research on ASER demonstrates that its design mediates major evaluation challenges in India, but empirical research must be done to understand how these design choices translate to new environments and the wider social and political consequences of these M&E translations.

Future Research

Throughout this paper, I have highlighted areas that are worthy of further investigation through empirical examples either from ASER or other evaluations using the ASER model. One example is examining whether the actual influences of ASER's evaluation align with its "model pathways" for results-based influence and process-based influence. Such research can further illuminate the relationship between an evaluation's design (emanating from its theory of change) and its potential for actual influence. Another beneficial avenue for future research is in examining the levels (e.g., global, national, district, community, and individual) at which

influence happens. It is both necessary and possible to further investigate levels of influence in a large-scale evaluation of social systems such as ASER because a defining feature of its design is that it operates hierarchically (i.e., in a vertical manner across many participants and stakeholders). Research on the levels at which ASER appears to be having influence should also analyze the manifestations or forms of the influence (e.g., policies, funding, deliberation, action, and norms). Investigating the forms and levels of ASER's influence can illuminate how close empirical examples of an evaluation's influence (from India) align with theories of evaluation influence, such as Henry and Mark's (2003) framework for evaluation influence conceptualized at the individual, interpersonal, and collective levels. Henry and Mark (2003) have identified several manifestations of evaluation influence such as "agenda setting," which signals when an evaluation outcome or issue "moves onto the media agenda, public agenda, or government agenda" (p. 303). Future analyses of ASER data may provide empirical proof of the forms of evaluation influence proposed by Henry and Mark, which is valuable given the limited research (i.e., Gildemyn's 2014 study) that investigates such frameworks for evaluation influence outside of Western countries.

CONCLUSION

When I first learned about ASER, I wondered, could a process like ASER help improve education in such a large, diverse democracy as India? Given that ASER is massive, I learned that figuring out how it works and if it works is quite complicated. The answers to those questions encompass issues of perspective and context, not solely data. Across India, there is nothing if not a diversity of perspectives and contexts. The large amounts of data that I collected during my 10-month ethnography of ASER capture some slice of that diversity of perspective in principally four contexts—Delhi, Rajasthan, Manipur and Tamil Nadu—and in each of the contexts, I asked people about their perceptions of education in India, of ASER, and of ASER's meaning and influence.

The three papers that comprise this dissertation tackle a narrower but important aspect of my initial question. The papers explore the methodological design choices, processes, and design influence of ASER as a means to begin (not finish) a systematic examination of ASER's potential for widespread influence on India's education system and on education globally. Nevertheless, the aspirations informing ASER are by its design and theory of change much broader than the improvement of India's education system. ASER aims for social betterment writ-large through shifting the culture in India around measurement, evaluation, and democratic participation (Goodnight, Paper 2). ASER is stratospheric in its aims, but the challenges the evaluation has surmounted thus far in its design and execution probably only foster lofty goals rather than deter them. The "making of history" (Goodnight, Paper 2, p. 71), going "in favor of rebellion" (Goodnight, Paper 1, p. 19), joining "this movement…in building a better India" (Appendix E)—they are grand, idealistic phrases that are simultaneously both the source and product of ASER's success. Whether one believes such statements or not is not simply a matter

of what data demonstrates or not; it is a matter of perspective and experience in relation to ASER, and I have attempted to make clear that subjective reality alongside my analysis of ASER's philosophy, goals, methodological features, processes, and data, which comprise its design.

APPENDICES

Appendix A: The Study Design and Paper 1 Analysis

Overview of Broader Study & Methodology

Interviews and observations of the entire ASER process (e.g., surveyor trainings, testing in households, and surveying in schools) were conducted within an ethnographic study spanning July 2014 to May 2015 with fieldwork in three states. This broader study examines ASER's design and investigates its influence on educational change in India. Employing a *vertical case study* approach, this research investigates ASER from its stakeholders' multiple perspectives to analyze ASER's operation and influence at various levels. According to Vavrus and Bartlett (2006), vertical case study methodology enables the comparison of "knowledge claims among actors with different social locations in an attempt to situate local action and interpretation within a broader cultural, historical, and political investigation," which allows researchers "to conduct research promoting full and thorough knowledge of multiple levels of comparison within a single vertically-bounded case" (p. 95). The narratives analyzed for Paper 1 explicitly reflect ASER from the social location of Pratham's leadership analyzed in context of ground- and policy-level dynamics to offer a comprehensive account of ASER's development.

Research Data & Analysis

Paper 1 is one piece of the broader ASER vertical case study. The paper's research questions are answered via semi-structured interviews with Drs. Banerji, Chavan and Wadhwa and analyzed via a narrative inquiry approach, which Rhodes and Brown (2005) explain has benefits in investigating organizations. Particularly, storytelling directly relates to the production of organizational case studies as "exemplar" stories for comparison and theory building (Dyer & Williams, 1991, as cited in Rhodes & Brown, 2005, p. 167). Addressing research questions via

the ASER architects' interviews required recognizing the data as intrinsically tied to their social location, thereby seeing their collective storytelling as constituting one among many stories that could be told about ASER. Consequently, analysis of their narratives was rooted in "a quest for meaning" that aided interpretation of ASER versus "a quest for scientific truth" (Rhodes & Brown, 2005, p. 167). Similarly, ethnography seeks to make transparent multiple interpretations of events or phenomena based on the knower's "location" (Gupta & Ferguson, 1997, p. 5). Narrative analysis aligns well with the "*interpretivist* world view" that frequently underlies qualitative and ethnographic research (Willis, 2007, p. 6).

In this paper, I was interested in the chronology of how things unfolded within each of the architect's accounts. This article being in large part about the evolution of ASER required paying attention to each narrative's linear qualities and the accumulation of events and decisions it revealed. After transcribing interviews, I read each one and wrote memos about salient themes, turning points (e.g., events), and connections that emerged, noting links between the architects' responses and background literature on India's educational policies and contexts. Afterward, proceeding transcript by transcript, I chunked pieces of discussion in each transcript through carving out dialogue by topic, event or decision. I moved pieces of discussion around if, for instance, during the course of the interview, the architect returned to a prior topic to make a connection or add clarification. To the extent possible, I also moved chunks of dialogue into a chronological order to offer the clearest representation of the linearity of events and decisions. After chunking dialogue and chronologically arranging each transcript, I highlighted parts that were relevant to the research questions and lightly edited these responses: removing repetitions, minor vocal habits like "you know," and fragments of dialogue that were tangents or reflected brief incoherent struggles with thoughts or language before articulating an idea. (These

omissions are indicated by ellipses). Next, I read each of the architects' chronological, edited narratives and analyzed the alignments and distinctions in

a) Their accounts of their participation in ASER's development,
b) Their perceptions of how ASER originated and unfolded,
c) Their understandings of ASER's goals and the implications of its design.

This analysis then enabled me to sketch a broader narrative of ASER's creation as it happened at different moments across multiple places with various actors. The sketch included major turning points and key decisions in ASER's development. I supported, verified, and refined this initial sketch of ASER by entering chunks of the architect's narratives into it and conducting a closer analysis of the narrative chunks (once placed within the sketch), which facilitated the identification of ASER's goals and implications of its design.

Appendix B: The Main Features of ASER

Feature	Description	Rationale
Basic Learning Assessments	One-page reading and arithmetic tests that each assess four levels of competency up to a grade-level 2 or 3 proficiency	The simplicity of ASER tests enhances volunteers' capacity to assess children's learning levels reliably. Tests are designed to straightforwardly assess foundational literacy and numeracy skills in a manner so even illiterate parents can decipher the tests' meaning.
Oral Testing	Learning assessments are administered orally, one-on-one by an ASER surveyor to all children (5-16 years old) in sample households	Oral testing increases the validity of ASER assessments because children, who have limited reading skills, are still encouraged to demonstrate what they "know how to do," unlike with pen-and-paper tests that rely on children's ability to read and understand directions in order to navigate and interpret test questions and tasks.
Household Approach	Learning assessments and surveys regarding children's school participation and home environment are collected in sample households	ASER samples children from households in villages (versus schools) and tests their learning abilities in their homes. ASER also includes household survey data ranging from children's enrollment status to parents' education levels. This approach allows for representation of children who are out of school and enrolled in various types of schools. It also enables collection of data on home factors that influence schooling (e.g., household assets and mother's education).
Government School Survey	Surveys of government primary schools in villages include indicators aligned to India's Right to Education Act (RTE) of 2009 that ensures equitable provision of primary education	ASER surveys on government primary school per village to assess its conditions and basic operation. Indicators on the government school survey are linked to goals expressed in RTE, such as number of multi-grade classrooms, availability of drinking water, accessibility of teaching and learning materials, and attendance levels of children and assigned teachers. Indicators are designed to produce findings for the public and government that illuminate the success of RTE implementation across rural India.
District-level Sampling	The sampling frame of ASER is designed to be representative at the district-level.	Because budget allocations and important decision-making happens at the district-level, ASER was designed to provide findings that were representative for districts (not blocks or villages). District representativeness has determined the number of villages, households, and schools included in the annual survey.
Volunteer Surveyors	Volunteers recruited from each district conduct data collection for ASER in villages.	After being recruited and trained, the volunteers are directed by ASER Master Trainers to travel to sample villages for village mapping, household surveying and testing, and school surveying. Using local volunteers to conduct the survey helps ASER to transcend vast cultural and linguistic differences across rural India and increases the trustworthiness of the data collected. Moreover, using volunteers minimizes costs and engages young people in educational issues and evaluation.
Partner Organizations	Partner organizations (e.g., government teacher training institutes, universities, and non-governmental organizations) help facilitate ASER at the district-level	The involvement of a partner institution for each district in every state also helps minimize the costs and logistical hurdles of conducting ASER for ASER Centre. Partner organizations provide personnel, space, and other forms of assistance that the Centre cannot itself maintain in the

		500+ districts that it conducts ASER. Moreover, partner organizations have vital local knowledge (regarding communities, language, culture, security issues, etc.) that makes possible executing ASER in districts with difficult terrains, social tensions, and other challenges.
"Easy-to-Digest" Findings	Tests and surveys are designed to capture simple, important indicators of children's learning and schooling as well as of government school quality	Tests and surveys request data that are relatively easy for volunteers to observe, collect, and tally. (e.g., Do children know the alphabet letters? Do schools have toilets?) They are also "easy-to-digest" (accessible) facts for education non-experts to analyze. This means ASER findings can be understood and debated by multiple stakeholders.
Annual, Public Reporting	The annual report is available free online and is disseminated to media, government officials, civil society organizations, and partners.	The report is released annually ahead of final national budget decisions so findings can inform decision-making and allocation of funds. ASER Centre purposefully publishes sample survey forms, tests, and a clear explanation of its sampling, data collection, and monitoring processes so the public can interpret its findings. Dissemination efforts include sending the report to members of parliament, the Ministry of Resource and Human Development (MHRD), the National Council on Education Research and Training (NCERT), state government officials, ASER partner organizations, and representatives from Indian and international civil society/policymaking bodies.
Independence	ASER is conducted annually independent of the Government of India (GOI).	ASER efforts are not funded by GOI and are not dictated in any way by central government. State governments in with District Institutes of Educational Training (DIETs) are the main ASER partner have chosen to support the ASER data collection effort by stipulating that their teacher training institutes provide volunteers and logistical assistance. Otherwise, ASER is conducted completely independent of the government, which prevents government interference with its data collection efforts or release of findings to the public.

Appendix C: ASER 2014 Partnership Agreement

ASER 2014
Roles and responsibilities of partner organisation and ASER Centre

Part I: Responsibilities of partner organisation

A. Human resource and district survey requirements:

- In each district, identify 65-70 volunteers (for one-weekend survey) or at least 35-40 volunteers (for two-weekend survey) to participate in the district training and survey.
- During training, ASER Centre staff and Master Trainers will decide whether the volunteers provided by the partner will qualify to do the survey.
- Facilitate and arrange a district-level training for a minimum of 15 hours over 2 days or 21 hours over 3 days. The training venue should be in a convenient location for the participants.
- Ensure that all participants attend all days of district-level training.
- Take responsibility that all volunteers following the survey procedures and rules.
- Ensure timely submission of all survey-related data and financial formats to ASER Centre/Master Trainers as per the instructions outlined in the Survey and Financial pack.
- Ensure payment is made to all the volunteers who have completed the survey as per agreed terms, and provide an undertaking of the same to ASER Centre.
- If survey or data from a particular village(s) is found to be of poor quality, then that village will have to be resurveyed by the partner organisation at no additional cost to ASER Centre.
- Complete honesty and integrity is expected from the partner organization.
- If there are any difficulties in getting the tasks completed, the difficulties must be shared with the Master Trainer and ASER Centre state team.
- One contact person should be assigned by the partner organization for each district to coordinate the entire district-level training of volunteers and support ASER Centre in completing the survey satisfactorily in the district.
- Partner organisation must ensure that the names of survey village or schools will not be shared with anybody other than the volunteers who will go for the survey. It is a usual practice in research projects to keep names of villages and households confidential to protect the privacy of respondents.
- If the partner will be providing Master Trainers,
    - Allocate 1 or 2 Master Trainers per district to conduct the district-level training of volunteers and monitor the survey (Refer to Roles & Responsibilities document for Master Trainers).
    - Ensure that the selected Master Trainers are present on all five days of a state level training at the pre-decided location.
    - Ensure the Master Trainers are not given any other responsibilities during the entire period of the survey.

B. Responsibilities of a volunteer:

- Attend all days and all sessions of district-level training.

- Be punctual and participate sincerely in the classroom sessions and field work during the training. If a volunteer's performance in the classroom, during field work or in the arithmetic and English tests is found to be unsatisfactory, the partner will be asked to replace the volunteer.
- Meet the desired level of performance in the quiz during the district-level training.
- Travel to the village assigned to him/her and conduct the survey as per the ASER guidelines.
- During the survey, a volunteer must:
    a. Visit the assigned village with an assigned partner and collect information as per ASER survey guidelines.
    b. Survey one government school with primary sections in the assigned village as per ASER survey guidelines.
    c. Select 20 households in the village using the 5th household rule.
    d. Survey and test children in each of these households as per ASER survey guidelines.
    e. Fill out all required formats carefully, legibly and accurately.
    f. Conduct the survey over a weekend (either Saturday & Sunday or Sunday & Monday).
    g. Conduct the survey with complete integrity and honesty.
    h. Be respectful of the villagers, families and homes of the respondents as well as teachers and children in the school.
    i. Share any difficulties faced during the survey with the Master Trainer.
    j. Will not share/reveal the names of the surveyed village and school with anyone.
    k. After the survey, submit the completed formats and survey booklets immediately after the completion of survey to the Master Trainer.

Part II: Responsibilities of ASER Centre

- Conduct five-day state-level training for all Master Trainers in the state and bear costs of the training as per the agreed norms.
- Bear the cost of lodging and food incurred by participants during district-level training and field work as per the rates and procedures as have been agreed upon in advance between partner organization and ASER Centre.
- Bear survey-related material and logistics cost as have been agreed upon in advance between the partner organization and ASER Centre.
- Closely monitor the survey in the district.
- Provide a Certificate of Participation to volunteers, Master Trainers and Partner Organisation for participation in the ASER survey.
- Share copy of the ASER 2014 report.
- Acknowledge the partner organisation in the Annual Status of Education Report -2014.

Appendix D: ASER 2014 Master Trainer's Description of Responsibilities

ASER 2014
Roles and responsibilities of ASER 2014 Master Trainers

Part I: Responsibilities of Master Trainers
Master Trainers will work full-time for ASER Centre during the ASER phase (as planned in the timeline). Depending on the timeline, this may be 1-2 months at a stretch. At least two Master Trainers will manage the survey in a district.

A. STATE-LEVEL TRAINING

- Attend all days and all sessions of ASER 2014 state-level training.
- Be punctual and participate sincerely in the classroom sessions and fieldwork during the state- level training. If a Master Trainer's performance in the classroom, during fieldwork or mock training session is found to be unsatisfactory, he/she will not be part of ASER survey.

B. DISTRICT-LEVEL TRAINING
Each Master Trainer will be allocated a district or more than one district. Usually two Master Trainers will be assigned to a district to conduct training and monitoring in that district. They will have the complete responsibility of ensuring that the survey is carried out honestly and in a timely fashion (as per the state phase-wise plans). Master trainers are expected to be in contact with the partner organization in these districts to set up and finalize all details related to the district level training and survey. Soon after ASER 2014 state-level training, a two-three-day district-level training will be conducted by Master Trainers with support from the ASER state team. In addition to making the necessary logistical arrangements for the district-level training, the trainers will also be responsible for imparting training to volunteers who will actually conduct the survey in that district. District-level training has three broad components:
1) classroom training sessions, 2) field visit to nearby villages to practice all tools and procedures 3) clarifications and feedback. If the Master Trainer(s) faces any difficulty in any aspect of their work in the district, they must contact the ASER state team immediately.

C. SURVEY
Master Trainers will be expected to:

- Lead and coordinate the survey in 30 villages in the district as per ASER survey guidelines.
- Be responsible for the overall survey quality in the district.
- Coordinate the survey with complete integrity and honesty.
- During the days of the actual survey, the Master Trainer is expected to be in the field monitoring the survey work and ensuring that the survey is being done with utmost accuracy and honesty.
- Speak to the Call Centre Coordinator and provide accurate information as requested by him/her.

- The names of the villages and of the households/children are confidential. The village list should not be shared with anyone other than the volunteers going to survey that village. THIS IS VERY IMPORTANT.
- Share any problems faced during the survey with the ASER state team immediately.
- Strictly follow survey procedures.
- Have complete information about the survey team's movement plan and track it on a daily basis.
- Assist the volunteers in conducting the survey if volunteers face problems.
- Be respectful of the villagers, their families, and homes of the respondents as well as of the teachers and children in the school that will be visited.
- Ensure timely submission of completely filled survey-related data and financial formats to ASER Centre as per the instructions given in the Survey and Financial pack.
- Counter sign all receipts, and ensure no mismatch between the amount paid to the volunteer and receiving provided to ASER Centre.
- Carry out all monitoring and recheck processes (phone, desk and field recheck) as per ASER survey guidelines.
- Conduct field rechecks, and resurvey in case of poor quality of data. Immediately contact the ASER state team in case of a poorly surveyed village and then follow the instructions that the state team gives them.
- The Master Trainer may not engage in any other responsibilities or work during the entire duration of the survey.
- Often districts are visited by ASER Centre state teams and central teams as part of the re-check process. The Master Trainer is expected to provide full cooperation during such re-check visits to the district.
- In case the ASER or state central teams find that the data is of poor quality, Master Trainers will be expected to manage resurvey in specific villages or in the entire district without additional pay.

Part II: Responsibilities of ASER Centre

- Conduct five-day state-level training for all Master Trainers in the state and bear costs of the training as per agreed norms.
- Bear the cost of lodging and food incurred by participants during district-level training and field work as per the rates and procedures as have been agreed upon in advance between Master Trainer and ASER Centre.
- Bear survey-related material and logistics cost as has been agreed in advance.
- Closely monitor and recheck the survey in the district.
- Provide a Certificate of Participation for participation in the ASER survey to the Master Trainer and to volunteers.

Appendix E: ASER 2014 Pledge

I, as a citizen of India, have decided to volunteer for the Annual Status of Education Report 2014.

This effort aims to engage citizens and the government in the process of ensuring quality education for the children of this country.

For the past 9 years, volunteers like me have travelled to the farthest districts and reached the remotest villages.

We have gone from Kashmir to Kerala and from Gujarat to Tripura, to conduct this survey and collect information with utmost sincerity.

Today, along with 30,000 volunteers across India, I am going to be a part of the largest citizen-led survey of India.

I pledge to carry forward the honest work and vision of those who have come before me and become an example for those who will follow me in the years to come.

I understand the importance of recording only correct information in the survey sheets and pledge to do so with complete honesty.

Under no circumstances, will I avoid my responsibilities during the course of the survey.

I pledge to uphold the integrity of the important role this movement has entrusted in me and play my part in building a better India.

Volunteers' Signature

1. ……. 2. ……..

* Note: Pledge is copied verbatim from the back cover of the ASER 2014 "Instruction Booklet" distributed at the ASER 2014 National Workshop in Aurangabad, India in August 1-6, 2014 (See ASER Centre, 2014a, p. 28).

Appendix F: Monitoring and Recheck Rates per ASER Guidelines

Field Monitoring Guidelines in a District with One-Weekend Period

	Day 1	Day 2	Total
Villages Surveyed (#)	30	(30)	30
Volunteers in District (#)			60
Pairs of Volunteers in District (#)			30
Pairs Called for Phone Monitoring (#)	30		30
Volunteers Field Monitored by MT 1 (#)	2	2	4
Volunteers Field Monitored by MT 2 (#)	2	2	4
Volunteers Field Monitored (#)	4	4	8
Volunteers Field Monitored (%)	7%	7%	13%
Villages Field Monitored (#)	2	2	4
Villages Field Monitored (%)	7%	7%	13%

The (30) indicates the 2nd day of data collection pertaining to the same villages and thus, not counted twice.
The above figures represent the minimums stipulated in the guidelines set forth in the training documents for ASER 2014 (see ASER Centre, 2014e).

[See next page]

Field Monitoring Guidelines for a District with Two-Weekend Period

	Wk1, Day 1	Wk1, Day 2	Wk2, Day 1	Wk2, Day 2	Total
Villages Surveyed (#)	15	(15)[a]	15	(15)[a]	30
Volunteers in District (#)					30
Pairs of Volunteers in District (#)					15
Pairs Called for Phone Monitoring (#)	15		15*[b]		15
Volunteers Field Monitored by MT 1 (#)	2	2	2	2	8
Volunteers Field Monitored by MT 2 (#)	2	2	2	2	8
Volunteers Field Monitored (#)	4	4	4	4	16
Volunteers Field Monitored (%)	13%	13%	13%	13%	53%
Villages Field Monitored (#)	2	2	2	2	8
Villages Field Monitored (%)	7%	7%	7%	7%	27%

Note: The above figures represent the minimums stipulated in the guidelines set forth in the training documents for ASER 2014 (see ASER Centre, 2014e). [a] The (15) indicates a second day of data collection pertaining to the same villages and thus, not counted twice. [b] The * indicates the same pairs of volunteers called a second time during a second weekend of survey in new villages and thus, not counted twice.

Recheck Guidelines for a District: One- and Two-Weekend Comparison

	One Weekend	Two Weekend
Villages Verified in Desk Recheck (#)	30	30
Villages Verified in Desk Recheck (%)	100%	100%
Households Verified in Phone Recheck (#)	240*[a]	240*
Households Verified in Phone Recheck (%)	40%*	40%*
Villages Verified in Field Recheck (#)	4	8
Villages Verified in Field Recheck (%)	13%	27%

[a] The asterisk indicates minimums set by ASER Centre that may increase if irregularities are detected

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. doi: 10.1007/s11125-015-9350-5

Alkin, M. (2013). Context-sensitive evaluation. In M. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed., pp. 283-292). Thousand Oaks, CA: SAGE Publications.

Alkin, M. (2011). *Evaluation essentials: From A to Z*. New York, NY: The Guilford Press.

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

ASER Centre. (2015). Annual status of education report (Rural) 2014 [Provisional]. New Delhi: ASER Centre. http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202014/fullaser2014mainreport_1.pdf

ASER Centre. (2014a). *Instruction booklet [Training document]*. New Delhi: ASER Centre.

ASER Centre. (2014b). *Frequently asked questions [Training document]*. New Delhi: ASER Centre.

ASER Centre. (2014c). *Letter for Sarpanch*. New Delhi: ASER Centre.Retrieved from http://img.asercentre.org/docs/Aser%20survey/Survey%20Process/communicationpacksarpanch_english_aser2014.pdf

ASER Centre. (2014d). *Quality control framework*. New Delhi: ASER Centre.Retrieved from http://img.asercentre.org/docs/Aser%20survey/Ensuring%20data%20quality/qualitycontrolframework.pdf

ASER Centre. (2014e). *Quality control framework: ASER 2014 monitoring and recheck guidelines [Training manual]*. New Delhi: ASER Centre.ASER Centre. (n.d.a). *Ensuring quality.* New Delhi: ASER Centre. Retrieved from http://aserblog.com/making-of-aser/ensuring-quality/

ASER Centre. (n.d.). *Sample design of rural ASER* methodology paper. New Delhi: ASER Centre.Retrieved from http://img.asercentre.org/docs/Aser%20survey/Sampling/Sample_Design_of_Rural_ASER_1.pdf

ASER Centre. (n.d.b) Vision and mission. New Delhi: ASER Centre. Retrieved from http://www.asercentre.org/About/Us/History/People/Partners/Support/etc/p/48.html

Banerjee, A., Cole, S., Duflo, E., & Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics, 122*(3), 1235-1264. doi: 10.1162/qjec.122.3.1235

Banerji, R. (2013). The birth of ASER. *Learning Curve, XX*(Section C), 85-88.

Banerji, R., & Duflo, E. (2015, January 22). Let's remake the classroom: Grouping children by learning level rather than age or grade could be a solution for poor outcomes., *The Indian Express*. Retrieved from http://indianexpress.com/article/opinion/columns/lets-remake-the-classroom/

Barrett, A. (2011). A millennium learning goal for education post‐2015: A question of outcomes or processes. *Comparative Education, 47*(1), 119-133. doi: 10.1080/03050068.2011.541682

Bartlett, L. & Vavrus, F. (2009). Introduction. In Vavrus & Bartlett (Eds.), *Critical approaches to comparative education: Vertical case studies from Africa, Europe, the Middle East, and the Americas* (pp. 1-18). New York, NY: Palgrave MacMillan.

Banerji, R., Bhattacharjea, S., & Wadhwa, W. (2013). The Annual Status of Education Report (ASER). *Research in Comparative and International Education, 8*(3), 387-396. doi: 10.2304/rcie.2013.8.3.387

Chouinard, J. (2013). The case for participatory evaluation in an era of accountability. *American Journal of Evaluation, 34*(2), 237-253. doi: 10.1177/1098214013478142

Chouinard, J. & Cousins, J. B. (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). Evaluation theory tree. In M. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 11-57), Thousand Oaks, CA: SAGE Publications.

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

Cousins, J. B. (2015). Book review: Evaluation essentials from A to Z. *American Journal of Evaluation, 36*(3), 417-419. doi: 10.1177/1098214015586830

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

De, A., Khera, R., Samson, M., & Kumar, A.K.S. (2011). *PROBE revisited: A Report on elementary education in India.* New Delhi.

Donaldson, S., & Picciotto, R. (Eds.). (2016). *Evaluation for an equitable society*. Charlotte, NC: Information Age Publishing.

Drèze, J., & Kingdon, G. (2001). School participation in rural India. *Review of Development Economics, 5*(1), 1-24.

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. doi: 10.1177/1098214014525257

Goodnight, M. (2017). 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,* 0(0), 1-19. [Published online January 18, 2017: http://www.tandfonline.com/doi/full/10.1080/03057925.2016.1266926]

Goodnight, M. & Bobde, S. (unpublished). Missing children in educational research: Comparing school-based and household-based assessments in the pursuit of education for all in India.

Government of India. (2009). The Right of Children to Free and Compulsory Education Act, 2009. (No. 35/ Registered No. DL-N04/0007/2033-9). New Delhi: Government of India Press.

Government of India. (2011). Census of India. Retrieved March 22, 2017 http://www.censusindia.gov.in/2011-Common/CensusData2011.html

Government of India. (2014a). Chapter 13: Human development. *Economic survey 2013-2014* (pp. 230-254). New Delhi: Ministry of Finance.

Government of India. (2014b, November 5). Planning commission: Functions. Retrieved June 2, 2016, from http://planningcommission.gov.in/aboutus/history/function.php?about=funcbody.htm

Govinda, R. (2008). Education for all in India: Assessing progress towards Dakar goals. *Prospects (00331538), 38*(3), 431-444. doi: 10.1007/s11125-009-9101-6

Govinda, R., & Bandyopadhyay, M. (2010). Social exclusion and school participation in India: Expanding access with equity. *Part of a special issue: Universal primary education in South Asia: A right that remains elusive, 40*(3), 337-354. doi: 10.1007/s11125-010-9160-8

Greene, J. (2016). Advancing equity: Cultivating an evaluation habit. In S. Donaldson & R. Picciotto (Eds.), *Evaluation for an equitable society* (pp. 49-66). Charlotte, NC: Information Age Publishing.

Gupta, A., & Ferguson, J. (1997). Discipline and practice: "The field" as site, method, and location in anthropology. In A. Gupta & J. Ferguson (Eds.), *Anthropological locations: Boundaries and grounds of a field science* (pp. 1-46). Berkeley, CA: University of California Press.

Hay, K. (2010). Evaluation field building in South Asia: Reflections, anecdotes, and questions. *American Journal of Evaluation*, 31(2), 222-231.

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

Hood, S., Hopson, R., & Kirkhart, K. (2015). Culturally responsive evaluation: Theory, practice, and future implications. In K. Newcomer, H. Hatry & J. Wholey (Eds.), *Handbook of practical program evaluation* (pp. 281-317). San Francisco, CA: John Wiley & Sons, Inc.

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

House, E. (2013). Work memoir: Ideas and influences. In M. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 198-207), Thousand Oaks, CA: SAGE Publications.

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

Jha, P., & Parvati, P. (2014). Assessing progress on universal elementary education in India. *Economic and Political Weekly, 49*(16), 44-51.

Kingdon, G. (2015, December 15). Indian schools are failing their students, *The New York Times*. Retrieved from https://www.nytimes.com/2015/12/16/opinion/indian-schools-are-failing-their-students.html

Kingdon, G. (2007). The progress of school education in India. *Oxford Review of Economic Policy, 23*(2), 168-195.

Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation,* 88, 5-23.

Learning Metrics Task Force. (2013). *Toward universal learning: Recommendations from the Learning Metrics Task Force (LMTF)*. Washington, DC: UNESCO Institute for Statistics and Center for Universal Education at Brookings Institution.

Mark, M., & Henry, G. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation, 10*(1), 35-57. doi: 10.1177/1356389004042326

Mathison, S. (2014). Research and evaluation: Intersections and divergence. In S. Brisolara, D. Seigart & S. SenGupta (Eds.), *Feminist evaluation and research* (pp. 42-58). New York, NY: Guilford Press.

Mertens, D. (2009). *Transformative research and evaluation*. New York, NY: The Guilford Press.

Morabito, S. (2002). Evaluator roles and strategies for expanding evaluation process influence. *The American Journal of Evaluation, 23*(3), 321-330. doi: http://dx.doi.org/10.1016/S1098-2140(02)00200-X

National University of Educational Planning and Administration. (2013). *DISE 2013-2014 flash statistics: Elementary education in India, progress towards UEE [Provisional]*. New Delhi: Ministry of Human Resource Development, Government of India.

Nawani, D. (2013). Reflections on annual status of educational research 2012. *Economic and Political Weekly, 48*(12), 19-22.

Nunneley, R., 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, NC: 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.

Pratham Education Foundation. (n.d.). *Concept note: Combined activities for maximised learning (CAMaL)*. Pratham Education Foundation. New Delhi. Retrieved from http://img.asercentre.org/docs/Capacity building/Kamaal ke Shikshak/camalconceptnote.pdf

Pratham Education Foundation. (2006). Annual status of education report (Rural) 2005 [Final]. Mumbai: Pratham Resource Center.

Pratham Education Foundation. (2016). About Pratham. Retrieved April 29, 2016, from http://www.pratham.org/about-us/about-pratham

Pritchett, L. (2013). *The rebirth of education: Schooling ain't learning*. Baltimore, MD: Brookings Institution Press.

PROBE Team. (1999). *PROBE: Public report on basic education in India [Centre for Development Economics]*. New Delhi: Oxford University Press.

Results for Development Institute. (2015). *Bringing learning to light: The role of citizen-led assessments in shifting the education agenda.* Washington, DC: Results for Development Institute.

Rhodes, C., & Brown, A. D. (2005). Narrative, organizations and research. *International Journal of Management Reviews, 7*(3), 167-188.

Rose, P. (2015). Three lessons for educational quality in post-2015 goals and targets: Clarity, measurability and equity. *International Journal of Educational Development, 40*, 289-296. doi: http://dx.doi.org/10.1016/j.ijedudev.2014.11.006

Rosenberg, T. (2014, October 23). In India, revealing the children left behind. *The New York Times*. Retrieved from https://opinionator.blogs.nytimes.com/2014/10/23/in-india-revealing-the-children-left-behind/?_r=0

Ryan, K., & Cousins, J. B. (2009). Introduction. In K. E. Ryan & J. B. Cousins (Eds.), *The SAGE international handbook of educational evaluation* (pp. ix-xvii). Thousand Oaks, CA: SAGE Publications.

Ryan, K., & Feller, I. (2009). Evaluation, accountability, and performance measurement in national systems: Trends, methods, and issues. In K. Ryan & J. B. Cousins (Eds.), *The SAGE international handbook of educational evaluation* (pp. 171-189). Thousand Oaks, CA: SAGE Publications.

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, NC: Information Age Publishing.

United Nations. (2016). Sustainable Development Goal 4. *Sustainable Development Knowledge Platform.* Retrieved from https://sustainabledevelopment.un.org/sdg4

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. doi: 10.2304/rcie.2013.8.3.279