RAND CORPORATION

Research Report

BENJAMIN K. MASTER, ELAINE LIN WANG, BRIAN PHILLIPS, REBECCA L. WOLFE, HAROLD D. GREEN

The BHP Foundation's Education Equity Program

Midterm Evaluation Findings (2018–2022)

For more information on this publication, visit **www.rand.org/t/RRA239-6**.

About RAND

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest. To learn more about RAND, visit www.rand.org.

Research Integrity

Our mission to help improve policy and decisionmaking through research and analysis is enabled through our core values of quality and objectivity and our unwavering commitment to the highest level of integrity and ethical behavior. To help ensure our research and analysis are rigorous, objective, and nonpartisan, we subject our research publications to a robust and exacting quality-assurance process; avoid both the appearance and reality of financial and other conflicts of interest through staff training, project screening, and a policy of mandatory disclosure; and pursue transparency in our research engagements through our commitment to the open publication of our research findings and recommendations, disclosure of the source of funding of published research, and policies to ensure intellectual independence. For more information, visit www.rand.org/about/research-integrity.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Published by the RAND Corporation, Santa Monica, Calif.
© 2023 BHP Foundation
RAND® is a registered trademark.

Cover image composite design: Rick Penn-Kraus, (illustration of children) syntika/iStock/Getty Images

About This Report

This midterm evaluation report is intended to summarize the progress to date of the BHP Foundation's Education Equity (EE) program, identify lessons learned, and recommend potential ways to strengthen the program going forward. The report also presents an update on the next steps currently being considered by program leaders and outlines the potential implications of lessons learned to date for the broader field. The primary audiences for this report include the BHP Foundation and its EE partners, as well as other actors who may be engaged in similar evidence-to-practice initiatives worldwide. This report is based on research funded by the BHP Foundation as part of an overall study of its EE program investment. The findings and conclusions presented are those of the authors and do not necessarily reflect positions or policies of the BHP Foundation.

RAND Education and Labor

This study was undertaken by RAND Education and Labor, a division of the RAND Corporation that conducts research on early childhood through postsecondary education programs, workforce development, and programs and policies affecting workers, entrepreneurship, and financial literacy and decisionmaking. This report is based on research funded by the BHP Foundation. The findings and conclusions presented are those of the authors and do not necessarily reflect positions or policies of the BHP Foundation.

More information about RAND can be found at www.rand.org. Questions about this report should be directed to Benjamin Master (bmaster@rand.org), and questions about RAND Education and Labor should be directed to educationandlabor@rand.org

Acknowledgments

We are grateful to the BHP Foundation for supporting the evaluation and this report. We especially thank Faith Rose, the Program Director for Education Equity for the BHP Foundation; Brodie Vansleve, the former Program Director; Muza Gondwe, Program Officer at the BHP Foundation; and James Ensor, Chief Executive Officer of the BHP Foundation, for their guidance. We also wish to thank each of the partner representatives who generously shared their time and insights with us as part of this evaluation. We thank Julie Bélanger and Julia Kaufman for helpful feedback that greatly improved this report. We also thank Samantha Bennett for her editorial expertise and Monette Velasco for overseeing the publication process for this report. We greatly appreciate the support of Kristin Leuschner, who assisted with communications analysis. The authors take full responsibility for any errors in the report.

Summary

This report presents midterm findings from the RAND Corporation's independent evaluation of the BHP Foundation's Education Equity program (hereafter referred to as *the program*). The program recognizes that education is a basic right and essential to ensuring that individuals can earn a decent living and participate as citizens; yet, around the world, children and underprivileged populations lack quality learning opportunities. Launched in 2017, the program has invested in a wide variety of activities aimed at addressing the need to better leverage evidence and data to inform educational practice and to help facilitate transformative improvements in the quality of educational delivery. In its first funding window, the program has supported five partner organizations (*partners*) whose efforts span diverse country contexts and target learners both within and outside formal educational systems. Partners' projects are intended to be mutually reinforcing in their orientation toward using evidence to enhance educational practice and policy, but they also vary widely in how they go about that work.

The RAND team's evaluation is designed to assess the program's impacts, share formative insights to inform its continuous improvement, and provide recommendations related to its strategy and theory of change. The evaluation is also intended to inform the field as a whole, both by surfacing key lessons learned from the funded work and by sharing an example of a philanthropic program's evolution over time in response to assessments of program progress and changing external conditions. RAND published the first report on this program in 2021 (Master et al., 2021). That report described the BHP Foundation's strategic vision at the time of program launch, summarized the partners and their projects, and presented RAND's approach to evaluating the program. In tandem with this midterm report, RAND is publishing several stand-alone case studies that dig deeper into specific strands of the program's work. A final summative report referencing all of the evaluation findings is planned for fall 2024.

Our analyses of program impacts are organized around the program's initially established goals and theory of change, though we recognize that the program strategy continues to evolve. Our methods include reviewing partners' reports, assessing partners' prespecified indicators of progress relative to inaugural goals, interviewing partner representatives, reviewing program-sponsored independent evaluations of partners' progress, and conducting case-study investigations of specific strands of each partner's work. Given the methods we employ, our evaluation is largely limited to analysis and interpretation of existing data and evidence. The supplemental data we collect are exploratory in nature, designed to help interpret that existing evidence rather than to directly infer causation about overall program effects. Our results should also be interpreted with consideration of disruptions caused by the global coronavirus disease 2019 (COVID-19) pandemic, which delayed or forced adaptations to some of the school- and place-based activities initially envisioned by program partners.

Key Findings on Program Progress

Informed by our analysis of prespecified key indicators of program progress, as well as by our case studies, partner interviews, and review of partner reports and independent evaluations to date, we note the following key findings related to the program's progress and impact:

- Program investments have resulted in meaningful enhancements to the body of evidence and tools available to decisionmakers seeking guidance on how to adopt and scale effective educational practices. They have also specifically contributed to a substantial scale-up of UN Women's Second Chance

Education (SCE) program worldwide and to the expansion of various practitioner and decisionmaker networks.
- Partners have seen the clearest evidence of progress with respect to, and have converged on promising practices around, evidence curation and dissemination. Such practices are grounded in providing timely, embedded, and demand-driven insights and recommendations to educational and policy leaders.
- However, widespread mobilization of evidence to affect decisionmaking in specific local systems, advocacy to successfully influence policymakers, and purposeful leveraging of the potential of partners' networks to facilitate improvements in educational practice remain nascent and, in some cases, have fallen short of initial forecasts. This suggests that the systemic reforms being championed by the program at a minimum may take more time to reach fruition and may prove more difficult to achieve than initially envisioned.
- In some instances, our ability to gauge the depth and significance of partners' achievements has been limited by the quality of the available indicators and their orientation toward partners' own implementation efforts rather than changes in audiences' behaviors.
- While partners have sought to address issues of equity in various ways, the program's overarching equity agenda remains unclear.

Opportunities to Strengthen Partners' Efforts

In addition to summarizing program progress to date, our analyses have also identified some key formative insights and associated opportunities to strengthen or amplify program-funded work going forward. We highlight the following considerations:

- Partners' significant progress in developing approaches to curating and disseminating evidence presents opportunities to (more) formally document and share insights about what has worked or not in these areas, both with each other and with the field as a whole.
- Promoting evidence-based reforms in diverse—particularly non-Western—contexts has, in many cases, proven difficult or taken longer than initially expected. Consideration of different audiences' demand for the types of insights, recommendations, and tools that result from the funded work will be important moving forward.
- Partners have strengthened and increasingly highlighted the importance of their networks to promoting educational improvement, but they would benefit from being more explicit and intentional in how they are leveraging those networks to achieve their goals.
- Examples of both top-down and bottom-up approaches to advocating educational reforms appear promising, but effects of advocacy are difficult to measure. Moving forward, it will be helpful to ensure that partners have clearly articulated theories and/or proofs of concept for how their advocacy work is intended to play out (e.g., how top-down approaches will reach local practitioners, or how bottom-up approaches will promote systemic change).
- Partners and the program have benefited from the program exercising flexibility and investing in partners' core work. At the same time, it may be helpful to consider guardrails around permissible pivots and ways to determine whether defining features and goals of the program are being met.

Implications for the Program

The program has invested in a wide range of approaches to evidence-to-practice work, with interim results that provide lessons for the future. In accordance with our evaluation results thus far, we offer the following recommendations for how the program's approach could be strengthened as it moves into the next phase of its work:

1. Clarify standards for what types of evidence the program considers sufficiently actionable and capable of improving learning outcomes.
2. Set clearer goals around equity, coupled with monitoring the portfolio's collective emphasis on addressing those goals.
3. Further articulate intended audiences and engagement models if the program aims to promote systemic change in particular local or national contexts.
4. Enhance the rigor of program evaluation activities by identifying targeted opportunities to assess the program's pathways toward impact.
5. Articulate a learning agenda tied to the program's hypothesized pathways to impact to help focus evaluation activities on those proof points that are most relevant to the program's strategy and to the field as a whole.

As program leaders plan for the next phase of their work, they are adapting their strategy in ways that are closely aligned with our findings and recommendations. In particular, the program has begun to implement a new approach to measurement, learning, and evaluation (MLE), with a goal of updating its theory of change to guide the next phase of its funded work and associated evaluation activities. Program leaders have also engaged in a landscape review of key trends in global education to help inform next steps.

Broader Implications of This Work

Our evaluation of the program's portfolio of investments also provides a useful reference point and takeaways for the field, including other organizations engaged in similar evidence-to-practice efforts.

- Program efforts have surfaced valuable insights regarding efforts to curate and disseminate evidence effectively. These include the importance of contextualizing evidence to increase transferability, solicitation of local perspectives to ensure the applicability of what is curated, and the need to balance promoting practices proven to work in one context versus related practices that, while not proven, may better fit a new context.
- The program's work has also highlighted the key role that organizational partners and networks play in ensuring the reach, relevance, and longevity of efforts to scale innovative programs and enhance evidence use in educational systems and among practitioners.
- The program's example highlights the challenges inherent in measuring the long-term impacts of evidence-to-practice work. Clear articulation of the envisioned pathways to impact and monitoring of near-term proof points to pressure test those hypotheses are needed to better identify which innovations can contribute most to improving educational outcomes.

Contents

About This Report iii
Summary v
Figures and Tables xi

CHAPTER 1
Overview of the Education Equity Program and the RAND Team's Evaluation 1
Program Theory of Change 2
About the Partners and Their Funded Projects 4
Overview of the RAND Team's Methods of Evaluation 5
Limitations 9

CHAPTER 2
Program Progress According to Key Indicators 11
What is the Extent of Program Progress? 11

CHAPTER 3
Formative Insights from Program Efforts 21
Partners Have Converged on Some Promising Practices for Curating and Disseminating Evidence-Based Practices, but the Evidence-to-Practice Gap Still Looms Large 22
Partners Have Had Success Establishing Networks to Support Their Goals, but Engaging with Network Partners Also Raised Unexpected Challenges 24
Both Top-Down and Bottom-Up Advocacy Approaches to Developing Enabling Environments for Evidence-Based Decisionmaking Appear Promising 26
Partners Have Capitalized on the Flexibility Afforded by the Program Investment in Their Core Activities 27
Partners Have Sought to Address Issues of Equity in Various Ways, but the Program's Overarching Equity Agenda Remains Unclear 28

CHAPTER 4
Implications of Our Findings for the Program and Emerging Next Steps 29
Recommendations for the Program Strategy 29
Emerging Next Steps for the Program 31

CHAPTER 5
Broader Implications of This Work and Next Steps for the Evaluation 33
Next Steps for the RAND Team's Evaluation 34

APPENDIX
Key Indicators Tables 35

Abbreviations 53
References 55

Figures and Tables

Figures

1.1. Education Equity Program Theory of Change (Fall 2018–Spring 2023) 3
1.2. Countries Where Partners Have Been or Are Engaged in BHP Foundation–Funded Project Work 6

Tables

1.1. Overview of Partners and Projects 5
1.2. Examples of Partner Activities Connected to the Program Theory of Change 7
2.1. Summary of Progress on Key Activity Indicators, by Activity Category 12
2.2. Summary of Progress on Key Interim Outcome Indicators, by Interim Outcome Category 16
3.1. Summary of Formative Insights and Considerations for the BHP Foundation 21
4.1. Key Categories of Knowledge Gaps Related to Partners' Efforts and Potential Methodologies for Addressing Them 32
A.1. Key Activity Indicators: Curate Evidence About Improving Education 36
A.2. Key Activity Indicators: Disseminate Evidence-Based Strategies 38
A.3. Key Activity Indicators: Develop Capacity to Leverage Evidence 41
A.4. Key Activity Indicators: Advocate to Develop Enabling Environments 43
A.5. Key Activity Indicators: Test Potentially Transformative Innovations 44
A.6. Key Interim Outcome Indicators: Increased Availability of Data and Evidence 44
A.7. Key Interim Outcome Indicators: More-Informed Decisionmaking 46
A.8. Key Interim Outcome Indicators: Networks of Purposeful Collaboration 48
A.9. Key Interim Outcome Indicators: Empowered Thought Leaders and Local Actors 49

CHAPTER 1

Overview of the Education Equity Program and the RAND Team's Evaluation

Education is a basic right and is widely recognized as essential to ensuring that individuals can earn a decent living and participate as citizens. An educated populace is also critical to a country's economic and social development (Roser and Ortiz-Ospina, 2016).

Yet, around the world, hundreds of millions of youth lack basic literacy and numeracy skills and quality learning opportunities to develop those skills. Children in low- and middle-income countries are disproportionately affected. A recent census shows that more than half of all children read below a minimally proficient level by the end of primary school; for low-income countries, the rate is 90 percent (World Bank, 2019). Moreover, secondary school enrollment rates have stagnated; in low-income countries, only about 40 percent of upper secondary school-aged children remain in school (United Nations Educational, Scientific, and Cultural Organization [UNESCO], Institute for Statistics, 2019). In many countries, gaps in educational outcomes for boys and girls are evident and troubling (Wetheridge, 2016). The coronavirus disease 2019 (COVID-19) pandemic has exacerbated the global learning crisis in many ways (World Bank, 2020). Children worldwide have suffered from school closures, limited access to remote learning, high dropout rates, and learning loss, not to mention attendant nutritional, health, and safety consequences. Given constraints in resources and infrastructure, among other factors, the impacts for children in low- and middle-income countries are particularly alarming (World Bank, UNESCO, and the UN Children's Fund [UNICEF], 2021).

The BHP Foundation, through its Education Equity program (hereafter, *the program*), has made global education one of its priorities for philanthropic investments. The program launched in 2017 and began to fund partner efforts in 2018, with a focus on leveraging evidence of effective educational programs and testing transformative innovations to enhance education and learning globally. These focus areas were informed by the foundation's analysis of the educational landscape and its assessment of how, as a new actor in the field, the foundation could best contribute to improving educational outcomes and equity. It identified as key areas for improvement the need to generate and make available more evidence and data about "what works" in educational practice, as well as the need for transformational—rather than incremental—improvements in how education can be delivered at scale. The process for how the program strategy was initially conceived and developed is described in more detail in the RAND Corporation's baseline report about the program (Master et al., 2021).

The foundation has sponsored RAND to conduct an independent evaluation to assess the program's impacts, share formative insights to inform its continuous improvement, and provide recommendations related to its strategy and theory of change. In this second published report about the program, we provide an update on key findings about impacts and insights from program efforts to date, summarizing the progress of partner activities that are still ongoing and are within the scope of the program's initial funding window. We also discuss the ongoing evolution of the program's strategy in light of lessons learned.

In our final public report, scheduled to be published in fall 2024, we will present summative findings about the program's implementation and impacts, drawing on all the work completed by partners during the

entirety of the initial funding window. Additionally, we are in the process of publishing several case studies on the work of the partners (Wolfe, Wang, and Master, 2023-a, -b, -c; Master et al., 2023). These case studies provide in-depth exploration of individual strands of partners' work, and they are referenced throughout this report.

Program Theory of Change

The program is investing in partner organizations and initiatives that can contribute toward improved access and equity in systems of education and learning globally. Through this program, the BHP Foundation has prioritized initiatives with the potential to facilitate what it describes as **transformative "step changes" in educational practices at scale** and to achieve these goals primarily by **using evidence and data to enhance the quality of decisionmaking, investments, and educational practice**. The intended beneficiaries of the program include students in traditional education settings, whose main goal may be academic achievement or degree attainment, as well as marginalized or underrepresented learners, including those outside the formal education system, who seek access to educational opportunities.

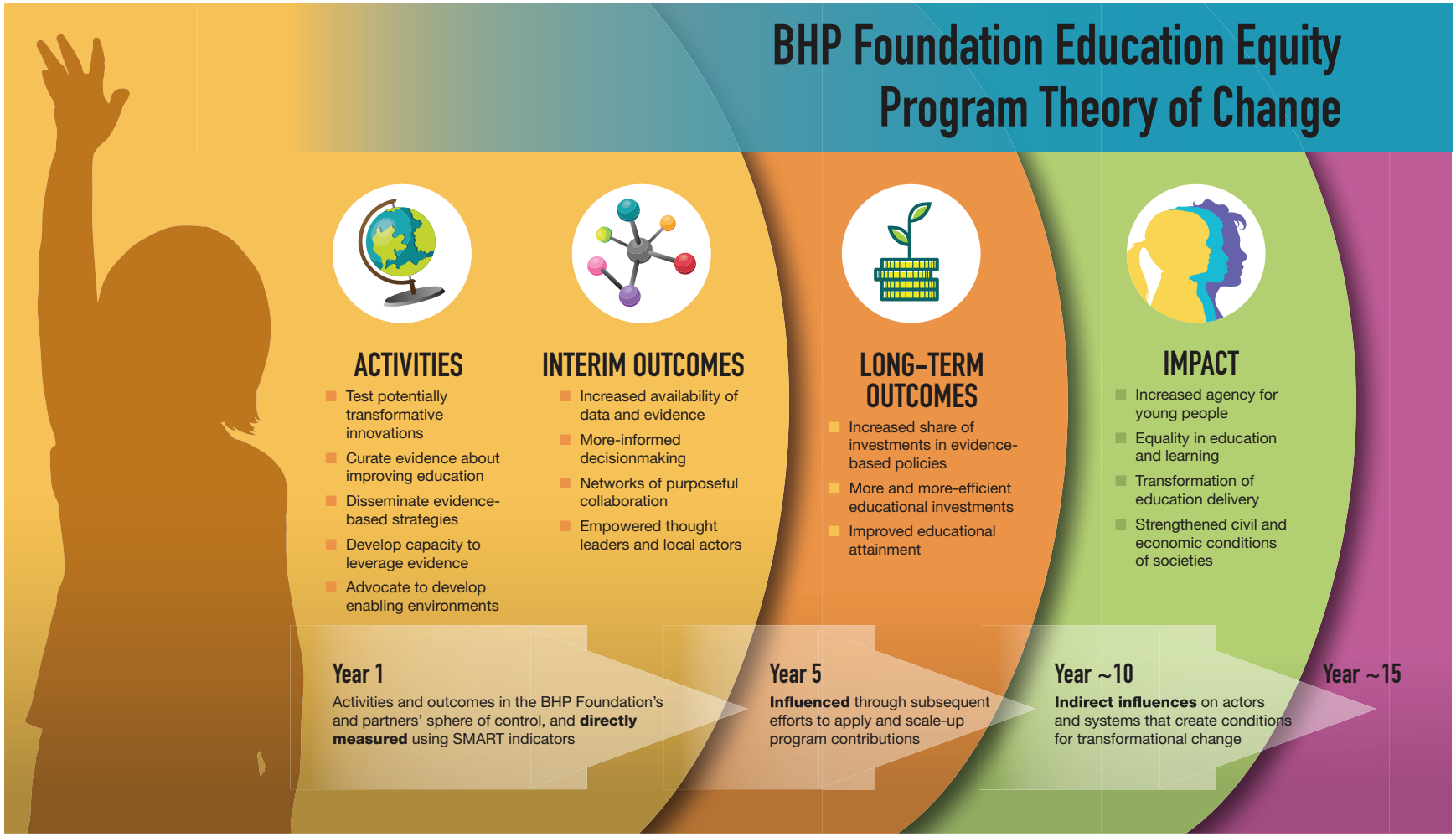
The program's initial theory of change, articulated by RAND researchers in spring 2020 and current as of spring 2023, is shown in Figure 1.1. The figure summarizes the **range of investments** made in the first phase of the program, as well as the **anticipated time frame for impacts**. It describes how the partners are collectively engaged in varied activities to enhance educational equity through the use of evidence. These activities include testing and generating evidence about new interventions or innovations, curating and disseminating (i.e., making publicly accessible) effective strategies, building capacity (e.g., skills and knowledge) of local decisionmakers to better leverage evidence about what works to improve educational outcomes, and advocating that societies and systems provide an enabling environment to facilitate more-informed investments in education and learning. The individual categories of activities are theorized to build on one another (e.g., better-curated evidence can, in turn, enhance dissemination efforts).

To date, the program has promoted efforts that refer to the various types of evidence that may be relevant to informing policy and practice. These include evidence from new or existing experimental and quasi-experimental studies about the causal impacts of interventions, observational and qualitative studies of unproven but promising interventions, expert and theory-based synthesis or interpretation of broad categories of prior research, practical knowledge and skills disseminated locally amongst collaborating practitioners, and pilot programs that—while untested—address unmet needs or key at-risk populations. However, on balance, the majority of the funded work references research about practices and policies whose causal impacts have been previously assessed.

The program's funded work is primarily focused on long-term and systemic change, with foundational investments in organizations and systems of evidence generation and use and relatively fewer investments in the direct provision of educational services. This **orientation toward systemic change** was part of the initial program strategy, with investments to develop tools, insights, networks, and proofs of concept that would demonstrate the utility of partner activities during the funded period but that were expected to require additional time and scale-up to achieve their intended impacts. Consequently, within the first five years of funded activity and the period of activity in scope for our program evaluation (i.e., 2018 through 2023), most outcomes are expected to be interim in nature. These include greater availability of evidence and data, empowered thought leaders engaged in more-informed decisionmaking, and the establishment of networks of purposeful collaboration. Long-term impacts, such as major effects on educational attainment or on the share or quality of educational investments, are expected to take more time to come to full fruition and are beyond the scope of our evaluation period.

FIGURE 1.1

Education Equity Program Theory of Change (Fall 2018–Spring 2023)



In most cases, the program funding also represents a significant portion of partners' total funding, and partners are provided significant discretion in allocating that funding. The program views its approach as contributing to the long-term organizational capacity of its partners, thereby providing both direct benefits from specific funded activities and indirect benefits from partners' broader contributions to global educational progress.

In preparation for the next phase of investments, in fall 2022, the program began a process of developing a revised articulation of its theory of change and pathways toward intended impacts. We discuss the progress of this and other ongoing program strategic planning in Chapter 4 of this report.

About the Partners and Their Funded Projects

The program has invested in a range of **global initiatives led by five established partner organizations** (referred to hereafter as *partners*), each of which brings its existing resources and networks to bear on its program-supported work. The partners are the Center for Universal Education at the Brookings Institution (CUE), the Education Endowment Foundation (EEF), the Global Business Coalition for Education (GBC-Ed), Teach for All (TFALL), and UN Women (UNW). Table 1.1 provides an overview of the partners and their funded projects, including the approximate scale and time frame of the funding provided by the BHP Foundation. Partly because of the COVID-19 pandemic and associated school closures, many of the partners' funding windows reflect extensions needed to allow them time to complete their work. In several cases, the partners are also pairing program funding with funding from other sources to support their project work.

While the partners are headquartered in the United States and Europe, their work is **international in scope**, with efforts potentially touched by the program in more than 100 countries worldwide. Figure 1.2 displays those countries where one or more partners are engaged in program-funded work.

Collectively then, the partners are engaged in a wide variety of work represented by the activities portion of the program's theory of change—**generating, curating, and disseminating evidence; building capacity; and advocating enabling environments.** Not all partners, however, are engaged in all of the activities listed—nor is that expected. Suffice it to note that the projects are intended to be mutually reinforcing in their orientation toward supporting the use of evidence about what works in education in policy and practice. Additional details about each partner's funded work can be found on the BHP Foundation's program website (BHP Foundation, undated).

While all contribute to the program theory of change, the **partners also vary widely in how they go about their work.** For instance, they curate and synthesize different types of evidence, from rigorous research to practitioner-generated insights; they target different audiences and stakeholders when developing capacity to leverage evidence; they engage in different types of advocacy efforts to develop enabling environments for the use of evidence; and they seek to address different inequities in educational opportunities and systems. While the partners' main workstreams are defined, the program has afforded them ample leeway to adapt their efforts as needed and has included as a program requirement that each partner engage in a midterm project evaluation and be responsive to its findings. As a result, partners' activities and focus could shift over time. Table 1.2 (updated from our first report; see Master et al., 2021) presents examples of partner activities, mapped onto the program theory of change.

TABLE 1.1
Overview of Partners and Projects

Partner	Type of Organization and Principal Location	Partner Project Name, Duration, and Funding Amount	Synopsis of Work
CUE	Think tank; United States	Millions Learning: Evidence and Innovation to Scale Impact; July 2018–June 2023 $8.0 million USD	CUE is working to identify effective practices and innovations, including around educational technology and family-school engagement, to accelerate learning; develop tools to facilitate the uptake and scale the impact of promising practices and innovations; and promote the use of data, including cost data, to drive decisionmaking and ultimately student outcomes.
EEF	Charity; United Kingdom	Building a Global Evidence Ecosystem for Teaching; April 2018–March 2025 $14.2 million USD	EEF is engaged in evidence synthesis, generation, and mobilization to inform decisionmaking among K-12 teachers, school leaders, and policymakers around school or classroom instructional practices to implement. EEF's key strand of work involves enhancing its existing toolkit of research evidence and developing contextualized versions of the toolkit for use outside the English context.
GBC-Ed	Nonprofit; United Kingdom and United States	Business Impact for Education Equity; January 2020–December 2023 $5.4 million USD	GBC-Ed is undertaking efforts to enhance the impact and effectiveness of corporate philanthropy. Its effort centers on developing and deploying a suite of resources, tools, and consultative services to businesses interested in making strategic investments in education.
TFALL	Nonprofit; United Kingdom and United States	Collective Leadership and Global Learning for Community Impact; October 2018–September 2023 $14.8 million USD	TFALL's global organization is working to accelerate the impact of the 60 partner organizations in its global network that collectively work toward developing locally rooted, globally informed leadership for system change. To accelerate progress, TFALL works to: (1) grow its footprint and scale (2) orient its network toward fostering the collective leadership of diverse alumni who work collaboratively at all levels of systems to transform education for the most marginalized students.
UNW	Multilateral; Switzerland and United States	Second Chance Education (SCE) and Vocational Learning Program; July 2018–December 2023 $25.0 million USD	UNW is supporting educational, employment, and entrepreneurial opportunities for marginalized women and working to reduce policy and societal barriers to such opportunities. Its core work centers on implementing a second-chance education program and engaging in advocacy efforts to change policies and social norms.

NOTE: USD = U.S. dollars.

Overview of the RAND Team's Methods of Evaluation

As detailed in our first published report (Master et al., 2021), we used data from a variety of sources to inform our evaluation of the impacts and implementation of the program. We systematically track a set of **key indicators** collected by each of the partners to monitor progress relative to targets. Moreover, we analyze **trends in partners' successes and challenges** and use qualitative data to help assess partners' progress more holistically. These data include semiannual **interviews** with partner representatives, review of **partners' reports** and of the **independent third-party evaluations** that the program mandates each partner participate in, and **case-study investigations** regarding specific strands of each partner's work and their distinct capabilities to support evidence-to-practice work. Additionally, in coordination with the BHP Foundation, we conduct ad hoc analyses of these and other publicly available data and research that can inform the program's strategic planning. As indicated in Table 1.1, most of the partners are entering the final year of their project funding; therefore, our assessment of their progress to date is not only the subject of this report but is also being used

FIGURE 1.2

Countries Where Partners Have Been or Are Engaged in BHP Foundation–Funded Project Work

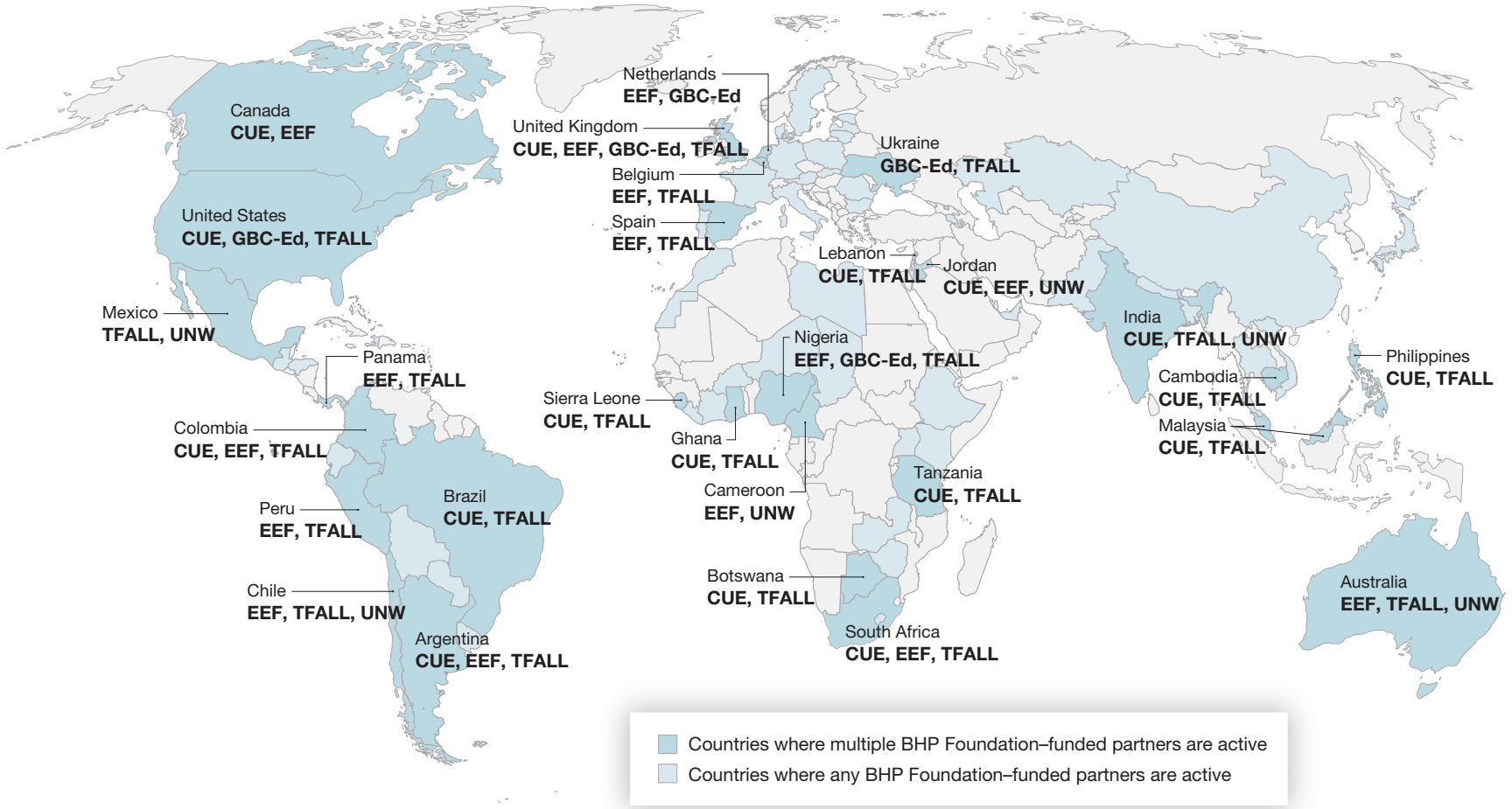


TABLE 1.2
Examples of Partner Activities Connected to the Program Theory of Change

Activities as Identified in Program Theory of Change	Examples of Partner Activities
Test potentially transformative innovations	• EEF has funded randomized controlled trials designed to test promising approaches in local jurisdictions. • TFALL has used the Community Impact Innovation Fund to test the effectiveness of a community's leadership development practices. • UNW has been gathering insights about what worked and did not work in its scale-up of Women's Empowerment Hubs and e-learning opportunities globally to support marginalized women in re-entering formal educational systems or accessing paid employment opportunities.
Curate evidence about improving education	• CUE conducted meta-research on specific topics (e.g., use of educational technology, family-school engagement practices) to identify approaches that support accelerated educational progress. CUE has distilled the key findings to produce playbooks to guide decisionmakers to take evidence-informed actions. • EEF's Teaching and Learning Toolkit summarizes evidence on how particular strategies have successfully or unsuccessfully influenced student achievement. EEF's global partners have been updating the Toolkit for local contexts and target groups.
Disseminate evidence-based strategies	• GBC-Ed has developed and begun to deploy resources that can help guide business and noncorporate actors toward investments that contribute to systemic change in the quality of education worldwide. • TFALL has generated and published insights about what differentiates high-performing and fast-improving communities through its Community Impact Lab.
Develop capacity to leverage evidence	• CUE has been supporting the capacity of its Scaling Labs partner institutions, training the Scaling Lab manager and staff to design, implement, and report on action research. • EEF has been building the capacity of a network of evidence brokers, providing training to these partner organizations on commissioning research trials and on the use of evidence. • GBC-Ed has been offering a suite of consultative services to support business leaders in making more-strategic investments in education. • TFALL has been supporting the ongoing growth of its global network of partners and developing alumni and partner leaders through its Global Learning events.
Advocate to develop enabling environments	• CUE has been supporting its Scaling Labs in reflecting on and identifying effective levers within the environment—in particular, political and technical aspects—that can contribute to scaling high-quality-learning educational innovations. • EEF has been supporting evidence brokers in playing an advocacy role in which it champions policy changes that help ensure high-quality learning experiences for all students in their jurisdiction. • GBC-Ed has been working to embed more-concrete education indexes in corporate Environmental, Social, and Governance (ESG) investment standards. • UNW has been raising awareness among local policymakers about the need to educate marginalized women and has been engaging in grassroots work to dismantle discriminatory social norms that pose barriers to this group of learners. • TFALL has engaged in global advocacy, including with the Education Commission, to establish collective leadership development as an important driver of educational transformation.

NOTE: The table provides illustrative examples. It is not intended to be a comprehensive inventory of all partner activities related to each category of activities in the program theory of change.

to inform next steps for the program. At the same time, the partners' final independent evaluations of impact have yet to be completed; therefore, our available data and the conclusions we draw remain preliminary at this stage.

Below, we briefly describe our methods of indicator analysis, followed by qualitative data collection and analysis. For more details, please see Master et al., 2021.

Indicator Analysis

Following suggested practice in program monitoring and evaluation (Bamberger and Hewitt, 1986; Haims et al., 2011; McDavid and Hawthorn, 2006), a core component of our evaluation involves tracking **prespeci-**

fied key indicators of progress toward program goals. We developed our indicator framework by compiling a library of more than 250 indicators being collected by the five partners as part of their own project reporting and applying SMART (specific, measurable, appropriate, reliable, and time-bound) criteria to select a subset of those to track on an ongoing basis. In essence, we selected tangible, objective indicators that we expect to observe during the limited time span of our evaluation. In selecting indicators, we prioritized those that clearly align with the activities and interim outcomes in the program's initial theory of change.

We designated each indicator as either an *activity indicator* or an *interim outcome indicator* but recognize that the boundaries between these steps of the theory of change can be fluid. We ended up with six to ten key indicators per partner, with a total of 23 activity indicators and 18 interim outcome indicators. Often, we considered the same indicator as a meaningful measure of several activities or interim outcomes. For example, a count of publications might document accomplishments in both dissemination and the underlying evidence curation work, while evidence of impacts on policy might indicate both more-informed decisionmaking and empowered leaders making those decisions. That there are more indicators for three of the activities in the program theory of change (i.e., curating evidence, disseminating evidence, and developing capacity) reflects the nature of the partners' work, as well as the aspects for which they are tracking progress.

Upon identifying the indicators, we worked with the partners to determine **target benchmarks for success** for their program work for the given indicators. In some cases, partners opted not to prespecify targets; this was usually because the nature of the work made a specific quantifiable target uninformative and potentially even counterproductive as an incentive to meet an arbitrary numeric goal. Therefore, some of our key indicators reflect specific quantitative measures, some encompass a set of closely related quantitative measures, and some are qualitative in nature.

Qualitative Data Analysis

In the qualitative component of our evaluation, we aimed to elicit details about partners' **project activities, progress, challenges, and lessons learned**. Our data sources included quarterly narrative reports that the partners prepared for the BHP Foundation, the report(s) from partners' independent midterm project evaluations, and 60-to-90-minute semiannual videoconference interviews with project leaders of each of the partner organizations beginning in spring 2020. In terms of the quarterly project reports and interviews, we focused for this report on those from spring 2021 to fall 2022. Finally, we drew on conversations and correspondence with program leadership to inform our understanding of the program's evolving theory of change and strategy.

For each round of partner interviews, we spoke with about 11 individuals across the five organizations (ranging from one to five participants per interview). In general, the interview protocol focused on partners' key strategies for achieving project goals. We asked about **progress on activities** aligned with the program theory of change, their project's **theory of change**, how the project is **addressing inequities**, and **challenges encountered and mitigating strategies**. We also clarified the indicators that they are tracking. Prior to each round of interviews, we reviewed the latest reports and updated the semistructured interview protocol to reflect the partners' recent progress. We customized the protocol to each partner to best probe its particular streams of work. We recorded the interviews and took detailed notes to facilitate review and analysis.

Analysis consisted of three core evaluation team members reviewing the reports and interview notes and debriefing as a team. Our approach most resembles reflexive thematic analysis (RTA; Braun and Clarke, 2006; Byrne, 2022), as we recognize researchers' role in interpreting patterns in the data and the importance of a collaborative approach in which multiple analysts reflect on the data together to achieve richer interpretations. In line with RTA, our process of coding and theme development is "flexible and organic" (Byrne, 2022), rather than reliant on a set of predefined themes and bound by a codebook. Our process of analysis largely followed the documented phases of RTA: familiarizing ourselves with data through (re-)reading, gen-

erating or brainstorming initial codes (i.e., labels for key ideas in the dataset), generating themes, reviewing themes, and defining and naming the themes. We converged on key learnings and emergent themes for each partner and across partners. We drew on established techniques to identify themes, including looking for repetition, similarities, and differences among data sources (Bernard, Wutich, and Ryan, 2016; Ryan and Bernard, 2003). To check the credibility of our findings, we triangulated across interviews and documents and sought confirmatory and disconfirming evidence (Denzin and Lincoln, 2003).

We are also conducting and publishing a series of **case studies on specific strands** of each partner's work and their distinct capabilities to support evidence-to-practice work (Wolfe, Wang, and Master, 2023-a, -b, -c; Master et al., 2023). Each case is intended to describe the arc and implementation of a specific partner initiative. The cases are designed to help demonstrate the theorized pathway between one or more categories of program activities (e.g., curating evidence, disseminating evidence-based strategies, developing capacity) and one or more interim outcomes (e.g., increased availability of evidence, more-informed decisionmaking, networks, and empowered local actors) as articulated in the program theory of change. From the set of in-depth cases, we aim to draw broad lessons for the partners, the BHP Foundation, and, potentially, the field at large. For each case study, we conducted two or more rounds of interviews or focus groups with partners' in-country collaborators. We also reviewed documents and artifacts relevant to the effort that the case study illustrates. We recorded and transcribed all interviews and transcripts. We coded the transcripts in Dedoose, a cross-platform internet application for mixed-method data analysis, identifying the strategies, key actors, challenges and facilitators, and preliminary outcomes involved in the case. We abstracted and analyzed our data by subtopics to identify themes. Representatives from the partner organization fact-checked each case.

Some of our case studies are more specifically focused on **networks of professional collaboration** that several of the partners view as key to their program-funded efforts. In these studies, we use extant records that provide evidence of individual or organizational affiliations or collaboration, as well as survey data from individuals about their professional contacts, to study the structure and types of interactions facilitated by the network. These case studies include additional quantitative and descriptive analyses of these network connections, as well as visualizations of the collaborative activity or contacts between individuals or organizations.

Limitations

By design, our methods have some key limitations. The quality of the indicators that we use to inform our judgment about areas of success and challenge for the program is outside our control. For the most part, they were not initially designed to test specific hypotheses about partners' intended pathways to impact and are often descriptive in nature rather than evaluative. Moreover, our own data collection—semiannual partner interviews and case study interviews—is exploratory; it is typically drawn from positive proof points of the partners' work. For example, given that the partners had input into the focus for each case study and the interviewees we had access to, the case narratives are likely positive or successful examples of partners' work. Moreover, given the brief time frame of each case study (i.e., data collection spanning one year), we are limited to documenting early signs of progress toward interim outcomes. Overall, our evaluation is not designed to infer causation about overall program effects. Instead, our **primary focus is on analysis of existing data and metrics** to interpret the available evidence in light of the program's strategy and objectives.

Organization of This Report

The remainder of this report is organized into four chapters. Chapter 2 outlines program progress to date, with attention to the key performance indicators collected about the partners' activities and impacts. Chapter 3 describes key formative insights generated by the partners' efforts, informed by our review of documents, interviews with partner representatives, and case studies of specific strands of their work. Chapter 4 presents recommendations that can inform the program's ongoing review of its strategy and learning agenda alongside updates from our interviews with program leaders about their planned next steps. Finally, in Chapter 5, we conclude with some reflections on the import of the program's example and of our findings for other global actors engaged in similar work.

CHAPTER 2

Program Progress According to Key Indicators

As detailed in the previous chapter, our evaluation of program progress combines analyses of prespecified key performance indicators (aligned with the program's initial theory of change) with qualitative analyses that can help to explain why and how certain funded activities have achieved success—as well as lessons learned from challenges. This chapter focuses on the analysis of indicators. We also present some implications and interpretations of these findings for the program.

What is the Extent of Program Progress?

In this chapter, we first present partners' progress in carrying out their planned program activities and then describe their progress toward achieving interim outcomes. We note that COVID-19 delayed or forced adaptations to some of the school- and place-based activities that program partners initially envisioned.

What Activities Did the Program Successfully Implement?

Table 2.1 summarizes the program's progress as measured by the 23 **key activity indicators** drawn from the five partners, according to the partners' reporting through the second or third quarter of 2022. The table is organized according to the activity or activities for which the indicator(s) measure progress. Then, within each bucket of activities, the text characterizes progress for each individual partner with relevant indicators (sometimes just one indicator, and, in other cases, multiple indicators, as noted in parentheses next to the partner's name). For example, the first row displays indicators related to testing potentially transformative innovations. EEF has one indicator that falls solely into this activity category. The second row describes a set of two indicators from UNW that relate to a combination of testing potentially transformative innovations and curating evidence. The last column highlights how progress to date compares with targets prespecified by the partners at the outset of their projects.

Tables in the appendix provide greater detail about each of the key activity indicators individually alongside their measured progress. The types of indicators vary across activities. For example, milestones in implementing trials or pilot programs are sometimes used to indicate progress toward the testing of innovations; milestones in the creation of new publicly available knowledge products speak to curating evidence; and publications and presentations speak to evidence dissemination. Capacity-building often relates to milestones for training or consulting with individuals, or the development of networks of practice. Finally, measures of advocacy activities include implementation of media or persuasion campaigns, as well as other activities to promote consideration of recommended practices by the public or decisionmakers.

To gauge program progress, we coded each indicator with one of six codes: *exceeding target* (meaning the partner has clearly exceeded the prespecified target); *meeting target*; *not yet meeting target* (the partner has not yet met a prespecified target, but the target is not out of reach); *falling short of target* (the partner is clearly falling short of a prespecified target and very unlikely to reach it); *difficult to assess relative to target*; or *no prespecified target* (see the "Guide to Key Indicators" box). Even as we report against the prespecified targets, we rec-

TABLE 2.1

Summary of Progress on Key Activity Indicators, by Activity Category

Activity Categories						
TI	CE	DE	DC	AD	Summary of Progress on Indicators	Progress Relative to Prespecified Targets
X					• **EEF (1)** has initiated ten trials across five jurisdictions to date, though progress was slowed by the pandemic and difficulties in obtaining matched funding; the geographic breadth of funded trials has improved recently.	• **Not yet meeting target** of two completed trials in each jurisdiction
X	X				• **UNW (2)** has developed more than 500 SCE training modules; more than 100,000 women have accessed SCE content.	• **Exceeding targets** for number of training modules (21) and number of women accessing SCE content (65,000)
X			X		• **UNW (1)** has collaborated with a network of partner organizations to establish 83 physical spaces or Women's Empowerment Hubs where SCE programs are available.	• **Exceeding target** of 44 physical spaces
	X				• **GBC-Ed (1)** interviewed 46 business leaders and 45 noncorporate entity leaders as part of its scoping, listening, and designing work to support tool development.	• **Exceeding targets** of 15–20 business and 15 noncorporate entity interviews
	X	X			• **CUE (3)** has produced about 160 knowledge products (e.g., reports, tools, blog posts, podcasts) across the project's three main workstreams, with the most products related to its Innovations strand, followed by Scaling and then Analyze Data.	• **No prespecified target**
					• **EEF (1)** released Toolkit 2.0 in September 2021 after coding more than 2,700 studies across the target of 34 strands; since then, EEF has begun automatic searching for additional studies for 13 strands to make Toolkit a "living review"; updating and integrating Early Years Toolkit into main toolkit also underway.	• **Meeting targets** for toolkit development and content
					• **TFALL (2)** is adding 50+ resources related to community impact (e.g., digital papers, case studies, snapshots, and artifacts) to its online resource bank per year. Resources highlighting partner practices in leadership development are also being developed.	• **Meeting target** of resources related to community impact • **Difficult to assess** whether resources highlighting partner practices align with initial vision
	X	X	X		• **TFALL (1)** held 30 learning experiences with 900 participants for partner staff, program alumni, and the broader education community in 2020–2021, and eight learning experiences with 1,300 participants through midyear 2021–2022.	• **Difficult to assess relative to target** because of evolution of types of learning experiences offered
	X		X		• **EEF (1)** has continued to expand the reach of its fellowship program, with 11 fellows to date across seven jurisdictions; role of fellows has evolved from a primary focus on evidence curation to developing contextualized approaches to evidence mobilization.	• **No prespecified target**

Table 2.1—Continued

Activity Categories						
TI	CE	DE	DC	AD	Summary of Progress on Indicators	Progress Relative to Prespecified Targets
		X	X		• **EEF (1)** has shifted its approach to evidence mobilization. Partners are now developing contextualized mobilization approaches, with several focused on teacher training. Pilot evidence hub networks have been established recently in two countries (Belgium and the Netherlands) that were not core jurisdictions for its funded work.	• **Falling short of target** because EEF has shifted its approach, deemphasizing replication of its evidence hub network model in each partner jurisdiction
					• **GBC-Ed (1)** has begun to roll out a set of tools and resources for businesses, including services at no cost, "Coalition" members, and several paid tiers of services.	• **Difficult to assess relative to target** because the tools and approach to piloting them have evolved over time
					• **TFALL (1)** activities to advance partner learning are ongoing, including development of an alumni strategy guide and 25 to 30 projects related to participant training, student impact, or alumni strategy in at least five regions.	• **Difficult to assess relative to target** because of project evolution.
		X	X	X	• **CUE (3)** initiated more than 750 consultations and held 120 events across its three workstreams to disseminate its research to audiences of policymakers and practitioners; recent highlights include events to launch its playbook for family-school engagement and its Tanzania and Côte d'Ivoire scaling reports, and data team consultations with more than 50 organizations to inform its work on costing, real-time data, and outcomes-based financing.	• **No prespecified target**
		X		X	• **UNW (1)** has supported its partners in engaging in 96 policy dialogues in individual countries during the pilot period**.** However, the organization has not yet met its target of supporting a global policy dialogue on women's right to education.	• **Exceeding target** of 26 policy dialogues in individual countries • **Not yet meeting target** for global policy dialogue
			X		• **TFALL (1)** has pivoted its monitoring and evaluation approach away from its initial vision of supporting partner organizations in participating in a community-level progress matrix and establishing local indicators for progress, thereby falling short of its target to have 50 communities develop progress indicators by 2021–2022.	• **Falling short of target** because TFALL has pivoted its monitoring and evaluation approach from initial vision
				X	• **EEF (1)** lists ten countries as recent successes in building "robust relationships" with policymakers and notes that all partners have "a strong working relationship" with policymakers in at least one country; these engagements are not clearly quantified, however.	• **No prespecified target**
					• **UNW (1)** and its partners have conducted 66 country-level awareness and advocacy campaigns related to changing social norms around SCE-related issues. However, they have not yet supported a global awareness and advocacy campaign.	• **Exceeding target** of 13 country-level campaigns • **Not yet meeting target** for global campaign

NOTE: TI = test potentially transformative innovations; CE = curate evidence about improving education; DE = disseminate evidence-based strategies; DC = develop capacity to leverage evidence; AD = advocate development of enabling environments. Numbers in parentheses indicate the number of key indicators for the partner for the category; includes data through the second or third quarter of 2022 (see the appendix tables for update date for each indicator) and assesses relative to initial key performance indicators even in cases where they have since been amended.

Guide to Key Indicators in Tables 2.1 and 2.2

Activity Categories

TI = test potentially transformative innovations
CE = curate evidence about improving education
DE = disseminate evidence-based strategies
DC = develop capacity to leverage evidence
AD = advocate development of enabling environments

Interim Outcome Categories

IA = increased availability of data and evidence
DM = more-informed decisionmaking
EL = empowered thought leaders and local actors
NW = networks of purposeful collaboration

Characterizing Progress Relative to Targets

Exceeding target = partner has clearly exceeded a prespecified target
Meeting target = partner has met a prespecified target
Not yet meeting target = partner has not yet met a prespecified target, but the target is not out of reach
Falling short of target = partner is clearly falling short of a prespecified target and very unlikely to reach it
Difficult to assess relative to target = nature of partner's work and reporting as it has evolved over time precludes characterizing progress relative to a prespecified target
No prespecified target = partner did not prespecify a target for the indicator

ognize that, in many cases, partners' activities have shifted relative to their initial plans (with the approval of the Foundation) as they responded to challenges and new opportunities. Of the 23 key indicators, only 15 had prespecified targets, and four of these 15 targets ceased to be relevant as partners' plans shifted over time. In all, assessing progress relative to the indicators is, in many respects, a qualitative exercise; the indicators are less quantifiable and measurable than the program likely initially intended them to be.

Despite the shortcomings we have noted, however, the indicators remain the best available measures of what partners have accomplished. Even in situations in which indicators have changed, they are still useful in illustrating shifts in strategy and the adaptive nature of partners' work. In all, the indicators are useful for interpreting the partners'—and the program's—progress with the proper caveats in mind.

Across all the activity indicators in Table 2.1, some of the clearest evidence of program progress in enacted activities according to partners' initial vision is in the areas of **curation of evidence (CE) and dissemination of evidence (DE)**. Collectively, the partners have largely met goals and/or achieved notable progress in preparing evidence for diverse audiences and engaging in a wide range of activities to share that evidence. Activities including tool and toolkit development, consultations, events, and publications have been implemented largely as initially intended. In this way, the program has contributed substantial new data and evidence to the field as a global public good. By and large, this information has been prepared for educational leaders and policymakers, and, while some elements are intended to be used by practitioners in the field, there is less evidence that practitioner-level dissemination has occurred at the initially intended scale.

Progress relative to intended activities in **testing potentially transformative innovations (TI)** has been mixed. Specifically, EEF has required more time than initially anticipated to establish its global trials for many reasons, including pandemic disruptions, difficulties finding partners equipped to carry out trials and evaluate them, and policymakers hesitant to invest time and resources in impact evaluation when "schools have no roofs" (Husain and Scandone, 2020). Separately, UNW has exceeded its targets for ambitiously piloting and testing its SCE learning modules in diverse contexts, generating substantial knowledge and experi-

ence about implementation, though these efforts did not initially include plans to evaluate the causal efficacy of the interventions.

Progress in efforts to **develop capacity to leverage evidence (DC)** has been difficult to assess relative to targets, largely due to lessons learned by partners over time and associated pivots in how they approached capacity development. Overall, multiple partners are still finding their way in this strand of work, as they have been challenged to tackle problems not previously addressed by their organizations or to tackle them in new contexts. EEF, for example, scaled back its initial ambitions for replicating national evidence hub models and has instead prioritized development of local capacity to use evidence through the work of fellows. TFALL has carried out an array of activities to build the capacity of its partners to develop and leverage evidence but has pivoted its monitoring and evaluation approach away from its initial vision of supporting partner organizations to develop community-level progress matrices. CUE's scaling efforts have identified key pain points on the road to scaling promising interventions in diverse contexts, and it has developed tools and consultative supports to help local actors address these barriers.

Lastly, activities entailing **advocacy of developing enabling environments (AD)** are ongoing and reflect a mix of targets achieved and activities where no target was prespecified. UNW has exceeded its targets both for in-country policy dialogues and for awareness and advocacy campaigns, though it has not yet supported global activities on these fronts. Meanwhile, both CUE and EEF have documented regular engagement with policymakers to advance their objectives. For CUE, this has come as part and parcel of its work to disseminate and encourage uptake of the evidence-based practices described in its research products, including through congressional testimony and memos to national-level policymakers in the United States. For EEF, this includes efforts to gain local policymaker buy-in around its multipronged strategy to test innovations, build a contextualized evidence base, and increase use of that evidence. Neither CUE nor EEF established prespecified targets or specific environments targeted for this work, and, in the case of EEF, only recently have partners begun to track and report on engagements with policymakers in a systematic way.

Next, we discuss what the available indicators can tell us about the interim outcomes that are the most immediate direct impacts stemming thus far from the program activities.

What Program Interim Outcomes Have Been Achieved?

All partners have begun to **demonstrate success in achieving one or more interim outcomes**, but our capacity to probe the depth and significance of these achievements is **somewhat limited by the quality of the indicators**. Table 2.2 displays a summary of progress toward the four interim outcome areas in the program's theory of change as measured by the 18 key indicators we are monitoring across the five partners. Compared with activity indicators, interim outcome indicators are less likely to be clearly quantified and more likely to lack prespecified targets (just eight have clear targets).

This sometimes stemmed from the emergent nature of the partners' work, wherein new workstreams that were not initially envisioned developed over time. More often, however, partners' interim outcome indicators could have reasonably included quantifiable targets but did not; this is a potential area for improvement in how the program sets goals moving forward. We hypothesize that partners may have not prespecified quantifiable targets, in some cases, partly to reduce the risk of not meeting a target; at the same time, they were also aware of the evolving nature of the work and may have wanted to allow for flexibility. Regardless, the import of the indicators that lack goals may still be assessed qualitatively.

Tables in the appendix list each indicator individually. Examples of indicators of **increased availability of evidence (and receipt of that evidence by audiences) (IA)** include such data as requests to receive new knowledge products. Indicators of **informed decisionmaking (DM)** include examples of the use of evidence to directly inform policies or practice. Indicators of empowered thought leaders and actors include engagements that enhanced their ability to reference or act on evidence. Finally, indicators of development of net-

TABLE 2.2

Summary of Progress on Key Interim Outcome Indicators, by Interim Outcome Category

Interim Outcome Categories				Summary of Progress on Indicators	Progress
IA	DM	EL	NW		
X				• **EEF (1)** has not yet published results from trials because of delays in conducting them; the lack of publications means EEF is not yet meeting its target regarding trial achieving its standards for evidence quality.	• **Not yet meeting target** for trials achieving standards for evidence quality
				• **TFALL (1)** reports "several high-profile mentions in external media and events" but has not yet developed standardized metrics for tracking the reach of stories and reports prepared for public dissemination.	• **No prespecified target**
X	X			• **EEF (1)** reports increased engagement with the Toolkit since version 2.0 was launched, with more than 850,000 page views from launch to March 2022 (metric not provided in more recent reporting) but has not documented through surveys whether there has been improvement in the rate at which school-based staff and system leadership are using Toolkit evidence to inform decisionmaking.	• **No prespecified target**
X		X		• **CUE (3)** has received more than 400 requests to share its work across its three workstreams, outstripping its capacity to do so; for example, the Innovations team fielded nearly 100 requests for guided workshops related to its family engagement work, and the data team has been invited to several events to speak about outcomes-based financing.	• **No prespecified target**
				• **UNW (1)** partners report that nearly 33,000 women have graduated the SCE program and are employed or self-employed, representing about two-thirds of graduates; however, inconsistencies in definitions, data collection, and reporting make it difficult to know the extent to which these outcomes would have occurred absent the program or to compare outcomes across countries and over time.	• **Exceeding target** of 17,500 employed or entrepreneur women among graduates
	X	X		• **EEF (1)** translated its engagement with policymakers into evidence-guided policy in the Netherlands, where the Toolkit was used to guide €5.8 billion in postpandemic school spending; ongoing efforts include a partner's engagement with four African governments to identify schools where evidence-based practices could be incorporated into teacher training.	• **No prespecified target**
				• **GBC-Ed (2)** has engaged 52 business leaders from 43 companies in its no-cost community, though the intensity of engagement with its tools is unclear; guidance material for noncorporate stakeholders has not yet been produced.	• **Exceeding target** of 10–15 businesses using its tools • **Not yet meeting target** of materials for noncorporate stakeholders
				• **TFALL (1)** surveys of learning experience participants show that 90 percent of attendees report the experiences as relevant and can apply knowledge or skills from the experience; the most-recent figures increased slightly from the prior year.	• **No prespecified target**
				• **UNW (1)** policy advocacy efforts have yielded 20 revised or in active review legislative frameworks or policies related to SCE across its pilot countries, while there have been two revised budgetary frameworks to date.	• **Exceeding target** of revised frameworks or policies • **Not yet meeting target** for revised budgetary frameworks

Table 2.2—Continued

Interim Outcome Categories				Summary of Progress on Indicators	Progress
IA	DM	EL	NW		
			X	• **EEF (1)** has not executed its plan to develop evidence hub networks in each partner jurisdiction, instead pivoting toward other approaches to evidence mobilization.	• **Falling short of target** of hubs with 50 engagements per year with schools
				• **UNW (2)** has engaged a network of more than 150 partners in supporting the development of its community learning spaces; UNW reports 178 peer networks of women at the country level; however, UNW has not established a global peer network.	• **Exceeding targets** for 70 partner engagement and 21 peer networks for women • **Not yet meeting target** for global peer network.
		X	X	• **EEF (1)** is seeking to consolidate and develop branding for its existing network of partners while cultivating relationships with several potential partners ranging from European countries (Netherlands, Belgium, Italy, France, Wales) to India and the Eastern Caribbean states and continuing to engage with a broader array of stakeholders.	• **No prespecified target**
				• **TFALL (2)** has grown its network to more than 88,000 program alumni and reports based on survey data that 70 percent are working or studying in fields related to its mission; the percentage of survey respondents reporting agreement with "mindset" questions has been around 80 percent for the past few years, while surveyed alumni and alumni response rates have varied.	• **No prespecified target**

NOTE: Numbers in parentheses indicate the number of key indicators for the partner for the category; includes data through the second or third quarter of 2022 (see the appendix tables for update date for each indicator) and assesses relative to initial key performance indicators even in cases where indicators have since been amended.

works include such examples as establishment of functional partnerships or connections among individuals that may enhance knowledge sharing and purposeful collaboration.

Partners' activities have resulted in IA, as measured by indicators that speak not only to the curation or dissemination of evidence but to some ***degree of engagement with that evidence*** by key audiences. However, the depth of engagement varies and has not yet been rigorously assessed by the partners or their independent evaluators.

EEF reports more than 850,000 views of its Toolkit 2.0, though it has not systematically tracked the rates at which practitioners across its jurisdictions are accessing this evidence base and using it to inform their decisions. Meanwhile, delays in carrying out trials have meant that trial publications have not yet made externally validated contributions to the global evidence base. TFALL notes that its stories and reports have received "several high-profile mentions" but has not yet developed metrics to clearly track the reach of its materials, and the materials themselves remain largely examples of promising practices from its work as opposed to rigorous impact studies or even qualitative studies that critically examine impact using something other than exemplary-case or convenience sampling methods. CUE reports an increased volume of externally initiated requests for its participation in meetings and events related to its three funded workstreams. CUE does not have prespecified targets for this metric, and monitoring what becomes of these engagements is a longer-term process not clearly captured by specific metrics.

Evidence of DM and **empowered thought leaders and local actors (EL)** is limited, reflecting both the nascent nature of much of this work and the difficulty that partners face in measuring the extent to which their efforts may be informing decisions or leading to changed behaviors.

Among the limited evidence available, there are nonetheless some notable examples of progress for these interim outcome categories. One example is EEF's successful engagement in the Netherlands, resulting in €5.8 billion in postpandemic school spending being guided by a menu of options drawn from its Toolkit evidence base. While this is a substantial achievement, we have limited information about the extent to which this has spurred use of Toolkit-recommended practices. Separately, UNW's five pilot-country partners reported a total of 20 legislative frameworks and two budgetary frameworks related to SCE at least being considered, the former exceeding and the latter falling short of prespecified targets. For example, the Government of India's National Institute of Open Schooling launched a new Inclusive Education policy in 2022 with technical inputs provided by UN Women in designing the policy. In this way, the organization may be empowering local political leaders to better address SCE issues.

Other indicators of more-informed decisionmaking or empowered actors are suggestive but similarly would require greater follow-through to validate. TFALL monitors whether participants in its learning experiences for partner staff and program alumni report that they benefited from the experience and are equipped to draw on the knowledge they gained in their work. These survey measures show consistently high rates of agreement with the questions posed (approximately 90 percent), but there is no specific target or reference point for their usefulness. TFALL plans, but has not yet completed, collection of follow-up information regarding the extent to which targeted audiences end up following through in applying any lessons learned. Though GBC-Ed's tools and resources for a business audience have just begun to be rolled out, it has exceeded its initial modest targets for the number of businesses engaging with its tools. The extent of that engagement and evidence that businesses are thinking differently about their investment strategies because of it remain to be seen.

Finally, UNW has surpassed its targets for the number of SCE program graduates employed or self-employed, though causal evidence about the extent to which its programs have been effective does not yet exist. Therefore, while the program has scaled successfully, it cannot be determined whether women who obtain employment or become entrepreneurs after exiting the program would have been motivated to do so even absent participation in SCE. Acknowledging this limitation, UNW has begun to expand its use of pre- and post- surveys and is exploring more-rigorous measurement to facilitate outcome-based financing of its efforts in some jurisdictions.

The final grouping of indicators contains those signaling whether partners have built **networks of purposeful collaboration (NW)**. Several partners report expansion in the number of individuals and organizations connected via their work in some capacity and have highlighted the increasing importance of those networks to accomplishing their goals. TFALL's global alumni network tops 88,000 people, up from 60,000 prior to the start of its BHP Foundation–funded work, and the organization reports, based on survey data, that 70 percent of its alumni are working in education or other fields related to its mission. A program-sponsored RAND case study of one such alumni network has highlighted that many alumni not only work in education but continue to engage in professional collaborations with each other (see Master et al., 2023).

UNW has exceeded its targets for the number of organizations collaborating with it to launch community learning spaces as well as the number of in-country peer networks of women, though, in the latter case, the definition of what counts as an in-country peer network is not clear. EEF is working to consolidate, brand, and create a shared online platform for its Evidence for Education Network, consisting of its current evidence brokers, and also is exploring the possibility of bringing additional brokers into formal partnership. However, difficulties in establishing evidence hubs in non-Western contexts has meant that EEF's development of local networks of collaboration around evidence use has not met targets. Finally, while CUE's pre-established indicators did not focus explicitly on network development, network development has proven to play a key supporting role in its efforts to share insights and inform decisionmaking.

What Are the Key Takeaways of the Indicator Analysis Findings?

On the whole, we observe that program investments have resulted in a **meaningful increase in and enhancements to the body of evidence and tools** available to educational leaders and policymakers seeking guidance on how to adopt and scale effective practices. The program has simultaneously helped to facilitate a **substantial scale-up of the SCE program** and contributed to the **expansion of practitioner and decisionmaker networks** of various kinds. However, **mobilization of evidence** in specific local systems, **advocacy efforts to influence policymakers**, and **purposeful leveraging of partners' networks** to facilitate innovations remain nascent, and, in some cases, these efforts have fallen short of initial forecasts. This pattern of results suggests the types of systemic reforms being championed by the BHP Foundation, at a minimum, would take more time to reach fruition and may also prove more difficult to achieve than initially envisioned.

In particular, efforts to promote evidence-based reforms in diverse—particularly non-Western—contexts have proven to be more difficult or to take longer than initially expected. For example, EEF's initial ambitions to replicate the UK evidence hub model in non-Western contexts have not yet been borne out. Similarly, CUE's involvement in promising efforts to scale interventions globally has yielded key insights regarding barriers and potential facilitators but has generated few instances of rapid scaling to date. This may be an important trend for the program—as well as other actors seeking to promote evidence-based reforms—to consider, moving forward. Partners have found that building trust, credibility, and influence in a new context takes deliberate effort and strategizing. Furthermore, there may be great value in placing in the first position the needs of the intended beneficiaries—reflecting on the extent to which intended audiences will have sufficient demand for the insights, recommendations, and tools associated with evidence-to-practice initiatives.

Additionally, while the **long-term impacts** of program contributions toward improved educational practices will, in many cases, be **diffuse and difficult to rigorously assess**, our indicator analysis highlights how the **program might have benefited from setting better grounded benchmarks** to monitor its desired interim outcomes. As we noted above, several interim outcome indicators could have usefully included quantifiable goals to clarify the intended scale or depth of impact but instead lacked prespecified targets tied to a specific theory about the scale of intended impacts. For example, the scale of consulting around corporate investments that was envisioned as a success by GBC-Ed or the degree of knowledge transfer anticipated by TFALL as a result of its global knowledge-sharing activities could have been better defined and measured. In future efforts, the program should consider taking a more active role in determining where additional quantifiable targets are appropriate and could help ensure that the ensuing results more clearly test partners' hypotheses for how their efforts will promote systemic improvements, on what timeline, and through which channels.

The program could also have benefited from setting targets or benchmarks more usefully oriented toward **documenting changes in practice and application** rather than partners' own implementation efforts. For example, if we wanted to know whether the program has increased availability of evidence, we can largely say this has been accomplished. However, if the goal was to know the frequency with which individuals accessing this evidence have used it to inform practice or policy, relatively few of the available indicators address this question. In future work, setting application-oriented benchmarks may also usefully challenge partners to critically assess the demand and readiness of their local partners and audiences to engage with the services and tools they are offering.

Informed by its results to date, the program has already begun to take steps to refine its approach to measurement, learning, and evaluation (MLE) going forward. We discuss the planned next steps in more detail in Chapter 4 of this report.

CHAPTER 3

Formative Insights from Program Efforts

In this section, we provide five key insights from our analysis of the evidence-to-practice work sponsored by the program. We drew on our review of the five partners' project reports and the reports of their independent evaluator, as well as our semiannual interviews with partner representatives and our case study investigations of specific strands of partners' work (Wolfe, Wang, and Master, 2023-a, -b, -c; Master et al., 2023). Each section below articulates a key insight. We include examples that support the insight, including vignettes in text boxes highlighting relevant case study investigations. We then provide attendant considerations for the BHP Foundation, with respect to either supporting current projects or planning the next phase of its funded work. It is our hope that the insights we present can inform the program's continuous improvement, and perhaps also the field of evidence-to-practice work as a whole. Table 3.1 summarizes the key insights and considerations discussed in this chapter.

TABLE 3.1
Summary of Formative Insights and Considerations for the BHP Foundation

Formative Insight	Considerations for the BHP Foundation
1. Partners have converged on some promising practices for curating and disseminating evidence-based practices, but the evidence-to-practice gap still looms large.	• Consider opportunities for partners and/or the program to (more) formally document and share insights about what worked or did not in curating and disseminating evidence. • Support partners to anticipate and meet the challenge of the evidence-to-practice gap, with greater focus on usability and audience needs.
2. Partners have had success establishing networks to support their goals, but engaging with network partners also raised unexpected challenges.	• Encourage partners to be more explicit and intentional in how they leverage networks (e.g., provide guidance on size or shape of partners' networks or support self-assessment of capacity to be responsive to their networks). • Provide additional supports and resources if funding new partners with less established partnerships to allow them to build networks to leverage in the work.
3. Both top-down and bottom-up advocacy approaches to developing enabling environments for evidence-based decisionmaking appear promising.	• Ensure partners have strong and clearly articulated theory and/or proofs of concept for how their advocacy work is intended to play out—clarifying, for example, how top-down approaches will ultimately influence local practices or how bottom-up approaches will contribute to systemic change.
4. Partners have capitalized on the flexibility afforded by the program investment in their core activities.	• If continuing to fund partners' core work, continue providing flexibility; at the same time, consider guardrails around permissible pivots and ways to determine whether defining program features and goals are being met.
5. Partners have sought to address issues of equity in various ways, but the program's overarching equity agenda remains unclear.	• Clearly articulate a definition of and agenda for equity. Clarify the extent to which partner efforts should be directed toward (specific kinds of) inequities. • Monitor the overall portfolio to ensure that sufficient collective emphasis is placed on remedying inequities as the program defines it.

Partners Have Converged on Some Promising Practices for Curating and Disseminating Evidence-Based Practices, but the Evidence-to-Practice Gap Still Looms Large

Consistent with the evidence from our formal indicators, our qualitative assessment highlights how program-funded work has converged on some **evidence curation and dissemination practices that are particularly promising**. In determining which evidence to champion to varied audiences, partners have tended to balance attention to extant research with factors such as its salience given current events, applicability to new contexts, and the demands and interests of the audience. For example, CUE prioritized family-school engagement given its timeliness—i.e., the surge of demand for help in this area during the COVID-19 pandemic. Moreover, CUE's resulting *Playbook for Family-School Engagement* ultimately showcased a combination of academic evidence and expert insights with promising practitioner strategies culled from diverse contexts (Wolfe, Wang, and Master, 2023b).

Similarly, in synthesizing evidence for its *Teaching and Learning Toolkit* in new contexts, EEF sought to balance rigorous evidence with the most-applicable evidence, opting to curate evidence-based recommendations alongside local examples that may have less rigorous evidence to support them but that illustrate how those recommendations might be particularly relevant in the given context (Wolfe, Wang, and Master, 2023c).

With respect to evidence dissemination, partners have all prioritized **"embedded" dissemination efforts** aimed at targeted groups of more-actively engaged practitioners, and more than one partner's midterm evaluation emphasized this as an area to focus on. Several partners have made substantial efforts to identify groups or contexts in which their recommendations are more likely to be picked up and used by decisionmakers or practitioners directly grappling with related challenges. *Co-creation* has also been a key theme; in this approach, decisionmakers work directly with the partners to customize evidence-based strategies to their own contexts. One partner representative described the values of co-creation as follows: "[Evidence-based] tools are one piece of a broader puzzle. The largest room for change is around the use of tools, not

Spotlight on EEF's Evidence Curation and Dissemination Approach

EEF's original UK-based *Teaching and Learning Toolkit* provides a summary of evidence on approaches to improving teaching and learning. The Toolkit presents information that teachers, school leaders, and policymakers can use to make decisions about which evidence-based practices to adopt, including the strength of supporting evidence, expected impact on student learning, and average implementation cost.

To ensure that the information will be relevant and actionable to its global partners—evidence brokers—EEF supported its partners in *contextualizing* the Toolkit. This involves identifying and appraising culturally relevant promising practices and summarizing them across each strand addressed in the Toolkit. The resulting summaries are accompanied by EEF's original summaries of evidence. The Toolkit is also translated into the local language.

Contextualization is intended to support interest in evidence and take-up of the Toolkit; it helps to counter skepticism that evidence generated primarily in Western schooling contexts is necessarily applicable elsewhere. EEF funded its partners to hire Global Fellows to conduct the technical work of contextualizing and disseminating the Toolkit. The fellows identify strategic leverage points specific to their context to disseminate the Toolkit. In Jordan, for example, EEF's partner, the Queen Rania Foundation (QRF), prioritized the Ministry of Education and teacher training programs as audiences. For more on contextualizing and disseminating the *Teaching and Learning Toolkit*, see the full case study focused on QRF's efforts (Wolfe, Wang, and Master, 2023a).

Spotlight on CUE's Evidence Curation and Dissemination Approach

CUE's approach to curating evidence typically involves reviewing available research alongside conducting original data collection. CUE co-created its *Playbook for Family-School Engagement* with the Family Engagement in Education Network (FEEN) in summer 2020. The FEEN consisted initially of 30 collaborating organizations from around the world that were already committed to family-school engagement and had an orientation toward evidence. By fall 2022, the network consisted of more than 100 members. On behalf of CUE, FEEN members gathered data from their jurisdictions, which CUE compiled and distilled into the published version of the Playbook. Specifically, FEEN members administered parent and teacher surveys, conducted focus groups and interviews with parents, teachers, and decisionmakers, and identified promising family-school engagement strategies. Subsequently, CUE used the FEEN to disseminate and champion the Playbook, with each partner tapping into its own network of districts, schools, and policymakers seeking to improve family-school engagement practices. CUE further supported mobilization of the Playbook by co-facilitating trainings or workshops with FEEN partners. For more on CUE's *Playbook for Family-School Engagement* efforts, see the full case study (Wolfe, Wang, and Master, 2023a).

about availability of tools. [Use] involves factors around the ecosystem. . . . Co-creation [helps] . . . get partners' and stakeholders' buy-in from the beginning on what the [tool] will be used for."

Despite partners' intentions and efforts, however, the evidence-to-practice gap is still a very real challenge. Both CUE and EEF, for example, realized that their published products—the *Playbook for Family-School Engagement* and the contextualized *Teaching and Learning Toolkits*, respectively—**needed to be made more accessible** for their intended end-users, including especially practitioners in schools and districts. For example, some organizations involved in the CUE's playbook work perceived the amount of information and the content in the playbook to be "overwhelming" even for school leaders; there were too many strategies to sift through and choose from. The process of (further) distilling the evidence base and strategies was largely unaccounted for in the partners' original plans, though the partners nimbly adapted. In response, CUE generated workshops and brief guidance materials to support implementation of the evidence-based strategies in educational practice. Following its independent midterm evaluation, TFALL worked toward articulating its previously implicit theory of how to support translating evidence into practice—that is, through a networked learning approach, in which local leaders and practitioners with common purpose share knowledge they gleaned from their on-the-ground work. From these localized examples, global insights (*evidence*) are generated that then influence local action (*practice*). TFALL is currently testing this model.

Turning to the implications of this insight for the program, a finding based on our indicators analysis suggests that the partners are delivering on the curation and dissemination aspects of their work; the qualitative insight here extends that finding by noting that partners are specifically most effective when motivated by more than existing research and theory and when **stakeholder input** is involved. There may be opportunities among partners and for the program to (more) formally document, share insights about, or make explicit what has worked and not in curating and disseminating evidence.

At the same time, given the evidence-to-practice gap, the program may need to help partners anticipate and meet the challenges—for example, through delineating this as a critical phase of the projects or supporting partners in planning their products with usability and ease of use for targeted audiences in mind. One related question for partners to consider is how to ensure that, in the process of making the products more usable and actionable for end users, the robustness of evidence that is curated and synthesized will not be diluted.

Partners Have Had Success Establishing Networks to Support Their Goals, but Engaging with Network Partners Also Raised Unexpected Challenges

As reported in the indicators section, partners have been generally successful in **building and/or expanding networks of collaboration**, and this continues to be core to their work. Partners intend for the networks not only to support scaling of products, interventions, and ideas but also to provide a shared space or community of practice that facilitates exchange of strategies and supports troubleshooting across network partners. For example, TFALL regards its global network as a foundation for generating and sharing insights across partners about how to cultivate collective leadership. Such learnings are shared through varied opportunities, including a global conference and virtual workshops and roundtables.

At the same time, because of their reach and keen interest in building and leveraging networks to support their system change efforts, some partners faced faster-than-expected growth in their network that challenged their ability to meet the resulting demands. In the first year of the project, for example, in consultation with the BHP Foundation, EEF decided to cease expansion of its Global Partner network, concentrating instead on strengthening and consolidating the existing partnerships "to maximize the chance of achieving the jurisdiction-specific goals" (interview with EEF project leaders).

Meanwhile, the FEEN that CUE established to help curate and disseminate evidence grew so quickly that CUE could not meet the demands of all organizations that had joined the FEEN. CUE could not provide all of them direct, individualized services and support to facilitate data collection or mobilization of Playbook strategies in their regions. To mitigate this challenge, CUE became more selective in the organizations it directly supported and helped build capacity of its larger partners. It also developed guided workshop materials that FEEN organizations could use to lead their own trainings on the Playbook. This challenge is detailed in the case study on CUE's *Playbook for Family-School Engagement* development (Wolfe, Wang, and Master, forthcoming-b). For both EEF and CUE, selectively supporting certain organizations is a reasonable

Spotlight on How TFALL's Alumni Network Connects Educational Organizations Across Peru

TFALL's core strategy involves developing networks of alumni in countries and regions around the world. Alumni work in a variety of sectors but share a common background in and passion for improving educational access and equity. TFALL believes that the presence of these actors at all levels of organizations involved in educational and policy improvement is a key ingredient to drive systemic improvements in educational practice and that this can occur through both their collective leadership and their ability to share knowledge and insights with each other and across the broader TFALL global network.

Using survey data from TFALL's Enseña Perú chapter, RAND conducted a descriptive analysis of alumni's professional networks in Peru. The data validate TFALL's assumption that participation in the program facilitates subsequent alumni professional collaborations that bridge the varied organizations where they work and additional organizations in their extended network. The collaboration network that we observed spans a wide variety of actors, including alumni and nonalumni working in educational organizations, nongovernmental organizations, and the private sector, with the largest hubs for alumni activity centered on alumni working in the Ministry of Education and in Enseña Perú itself. We also find that alumni are engaged in professional collaborations with nonalumni in organizations that leaders at Enseña Perú consider priority partners for local educational improvement. For more details, see the full case study (Master et al., 2023).

mitigating strategy given limited resources; however, this could result in selection bias, wherein the strongest organizations are preferred, and opportunities to develop or affect other interested organizations are missed.

Another challenge in relying on a vast network of partners to scale the intervention or ideas pertains to **delineating the defining or core aspects of the intervention** or idea from elements that are negotiable, which may be modified or adapted by the network partner. UNW's country-based SCE teams, for example, have used different models of program implementation. This has led UNW to reflect on how to consolidate them into one or a few program models that can be scaled while also tailored for different countries. A representative explained: "Globally, the biggest challenge is consolidating [an SCE] model that can be scaled up and can be tailored for different countries. We know what has and has not worked [in various countries]. We have identified some common challenges. The [difficulty] is in finding the agreement [while] not losing out on the knowledge gained from the country-specific implementation [of the program] . . . to standardize in a way, [but ensure it is] broadly applicable." EEF has also had to or will need to address this challenge, for example, in deciding which aspects of the Toolkit are essential to retain in the contextualization process and in making it friendly to end-users, as well as which elements of the original evidence-hub model are essential for their partners to replicate.

For the program, there are takeaways from both the partners' successes and challenges around establishing and leveraging networks. One observation, related to the partners' success in this area, is that the **program selected partners that were already well established** and had prior relationships with a large number of organizations and individuals that could be enlisted as collaborators; the partners are particularly well positioned to develop networks. If the program were to fund other organizations with less-established partnerships, partners would likely need more material resources and lead time to build networks that they can then leverage.

A second takeaway, given the challenges of responding to and leveraging networks that we documented above, is that the program might **encourage partners to be more intentional in the next phase about how to leverage their networks**. The program may provide direct guidance on the size or shape of partners' networks or support partners to self-assess their capacity to be responsive to the network that they develop and their use of networks. The network analysis featured in our case study of TFALL (Master et al., 2023) is an

Spotlight on UNW's Approach to Developing an Enabling Policy Environment to Support Second Chance Education and Employment Opportunities for Women

UN Women's policy change efforts typically involve working with local and national-level governments and government agencies with the goal of influencing them to identify and act on opportunities to support and sustain SCE-style solutions for women by writing, adopting, and implementing regulations, procedures, or laws or performing other administrative actions. Among other efforts, SCE India supported the National Institute of Open Schooling (NIOS) in revising its gender policy to ensure that more women and girls are able to participate. They also worked with the State of Maharashtra's Skill Development, Employment, and Entrepreneurship Department in creating and implementing new policy aimed at increasing entrepreneurial activities among women through the establishment of a permanent cell, the Women's Entrepreneurial Cell. UNW and SCE India's primary strategies included partnering with organizations already inclined to improve conditions for women and installing an SCE consultant as a liaison to the organization. This person engages in relationship-building, fosters open dialogues, and provides technical expertise. Furthermore, SCE India helped establish structures—such as a permanent gender policy steering committee within NIOS—to ensure that policy changes will be sustained. For more on UNW and SCE India's policy change efforts, see the full case study (Wolfe, Wang, and Master, forthcoming-a).

example of one approach that could support such reflection. For the next phase of work, the program should focus on what partners are doing with the networks, support them to be more intentional, and assess how effective the networks are for sharing knowledge or for facilitating purposeful collaboration.

Both Top-Down and Bottom-Up Advocacy Approaches to Developing Enabling Environments for Evidence-Based Decisionmaking Appear Promising

Partners' advocacy efforts to develop enabling environments for evidence-based decisionmaking reflect two different approaches, both of which appear to be promising. One is **top-down**. Several partners have placed bets on collaboration with system leaders to maximize their impact. Examples of this include EEF's partnership with the government's National Initiative for Education Research in the Netherlands around the use of their tool to facilitate COVID-19 response spending, and UNW's efforts to embed consultants in key ministries in India to enhance government-led efforts around gender inclusion.

Similarly, brokers in EEF's global network with ties to local education and teacher preparation systems have sought or are seeking to embed their Toolkit or Toolkit-recommended practices directly into education systems. For example, as documented in our case study (Wolfe, Wang, and Master, forthcoming-c), EEF's Jordanian partner, QRF, is working with other local partners to try to integrate its contextualized Teaching and Learning Toolkit into local universities' teacher preparation programs and possibly link participation in a course about the Toolkit to credentialing. In addition, the recent work by GBC-Ed to embed more-concrete education indexes in corporate ESG investment standards represents an effort to steer business financing by modifying the architecture in which investors evaluate organizations' impacts on these "externalities." One representative explained, "Within this nebulous world of ESG, there are metrics that companies are being rated on that are unrelated to corporate social responsibility goals but could be tied more closely to them. So we're exploring this, making all of this more streamlined. . . . We hope to create something that will help companies align themselves on ESG in a way that makes sense."

Other efforts to develop enabling environments for evidence-based decisionmaking are much more **bottom-up**. In particular, TFALL's alumni network represents a major bet on the importance of not just influencing highly placed decisionmakers but establishing a network of individuals at every level of the educational system who share a common mindset and commitment to educational improvement. In another example, UNW has made efforts to support marketing campaigns and community conversations that shift cultural norms around gender inclusion and build support for facilitating women's access to education and employment opportunities. In Cameroon, for example, UNW's implementing partner has been actively working with local religious leaders to host community meetings to discuss attitudes toward women's education and employment and practical steps to supporting women in these respects.

It is difficult to say whether one approach (top-down or bottom-up) is likely to be more productive than another, or indeed whether both may be necessary. The choice may depend on a range of factors, including partners' existing relationships with organizations. Partners consistently conveyed that the work of advocacy involves **building relationships and trust first**, which enables access to messengers and decisionmakers. Where the program intends to specifically influence high-level policymaking agencies, it may consider articulating this and selecting partners accordingly. Otherwise, the program can consider supporting partners in examining and articulating their strategies for advocacy efforts—in light of known and anticipated factors (e.g., existing relationships, context).

As indicated in our case study on UNW's advocacy work (Wolfe, Wang, and Master, forthcoming-a), the road from advocacy to policy change or shifts in attitudes, norms, and behaviors tends to be long, and the

effects of such efforts are especially difficult to measure. Given this, it is particularly important to ensure that partners have a strong and **clearly articulated theory and illustrative proofs of concept** for how their advocacy work is intended to play out. For example, how will a proposed top-down approach lead to change beyond a policy on paper? How will it lead to meaningful changes in attitudes, mindsets, and practices? Meanwhile, a bottom-up approach requires a theory about how a partner's efforts are expected to result in systematically improved conditions.

Partners Have Capitalized on the Flexibility Afforded by the Program Investment in Their Core Activities

As we found in a previous, unpublished interim report, the program's strategy is characterized not just by investments in individual project activities, but also by **support for partners' broader organizational missions**. This is reflected in the broad latitude that has been afforded to the partners to adapt their strategies and tactics to changing conditions over the period of funding, and in the formative emphasis of the midterm independent evaluations.

Partners have capitalized on this flexibility. In its third year, prompted by midterm evaluation findings, TFALL revised its project purpose and theory of change. TFALL articulated a clearer set of three learning activities and how they support the goal of developing and measuring impacts of collective leadership. In another example, GBC-Ed's new tiered business model and its work to integrate concrete education indexes in corporate ESG investments are also borne out of the organic nature of their work. These were made possible by the BHP Foundation's investment in the partner's core business development rather than strictly in a prespecified set of activities. A representative from one partner organization said, "I like [the BHP Foundation] approach as the grant agent. Their approach has let [our organization] work through an adaptive process rather than micromanaging the process. It's enabled us to think bigger rather than, 'we need to accomplish the things we put down [in the proposal].' So, flexibility to think bigger and evolve over time . . . [the foundation] is allowing us to grow with the project."

The pivots that the partners made in light of the COVID-19 pandemic also provide evidence of the potential **benefits of program flexibility**—namely, that it allowed partners to be responsive to pressing needs while continuing to make progress toward their primary goals and expanding their strategies in ways they might not have otherwise imagined. Specifically, in making necessary adjustments to the pandemic crisis, partners found opportunities to increase the relevance and urgency of their work. For example, the pandemic spurred a renewed CUE focus on and reception by audiences for evidence around family-school engagement strategies. Meanwhile, some UNW country partners shifted more heavily toward online learning. While this transition was a challenge at first, in many places, it led to strengthened capacities to operate online and increased access for women who might not have been able to access a physical SCE hub. In another example, several EEF global trials were delayed by pandemic-induced school closures. This led EEF to consider tailoring the trials to better address learning losses that children faced because of COVID-19.

The previous examples suggest that, to the extent that the program intends to fund partners' core work, providing flexibility may be key to supporting partners in thinking innovatively, improving their mission, and meeting the demands of the moment. One attendant consideration, however, is whether the program would need to provide guardrails about the kinds or size of pivots it would approve of and how it determines whether the fundamental intentions of the program are still being met. Or, if the program is less concerned about partners' pivots and the stability of program features, it should at least aim to articulate a clear set of shared goals and agreed-upon outcomes to help ensure that those targets remain in view.

Partners Have Sought to Address Issues of Equity in Various Ways, but the Program's Overarching Equity Agenda Remains Unclear

As its name suggests, the Education Equity program was initially conceived as having a major focus on enhancing educational equity. Correspondingly, each of the five program partners' activities are intended to reduce inequities in education and learning, either through direct service provision (such as UNW's SCE program activities) or, more often, through the promotion of evidence and best practices that could enhance educational system performance for underserved groups. Moreover, partners have made meaningful efforts to ensure that equity is a consideration in the process of their work. For example, most of the partners have intentionally sought to build networks in countries and/or communities that are underresourced and that demonstrate the greatest need. They have also been selective in the types of evidence-based practices that they have championed or curated to identify solutions relevant to those contexts. One partner representative explicitly said that in their project, they regard "equity as an end, with *end* meaning more-equitable outcomes for students in communities where [we are working]. . . . Also [equity as] a means to an end, [for example], in community partnerships—who has power to determine the outcome and how to pursue it."

However, the program overall has not clearly articulated the extent to which partner efforts should be **addressing inequities versus general educational improvement**, nor has it specified whether particular inequities should be central to the work. Indeed, one partner reflected, "I don't think we've had a formal conversation with [the BHP Foundation] about how they think about equity," while another partner characterized its own and the BHP Foundation's project focus on equity as "implied." Correspondingly, the partners have not systematically tracked (via indicators or otherwise) their collaborations or programmatic successes to date relative to any benchmarks that could help assess whether the program as a whole is sufficiently attuned to issues of equity or is making progress toward equity-related goals. Instead, each partner has attended independently to these issues, and it is difficult to assess whether the program is prioritizing enhancing issues of equity (defined either broadly or specifically) to the degree it initially set out to do. At least one partner agrees that both it and the BHP Foundation have more work to do to further clarify the emphasis on equity.

Program results to date also suggest a **potential tension** between the focus on evidence-to-practice work and priority on equity. By its nature, evidence is a global public good that is capable of "lifting all boats" and may tend to be **more accessible to better resourced countries and systems**. As EEF's midterm evaluator, NatCen, noted, "The complex context of low- and middle-income countries may mean that the education system becomes more divided, with better resourced schools and children from more privileged families benefitting from evidence-based interventions. To ensure equity, it may be necessary to find novel approaches to engage schools with the most disadvantaged students" (Husain and Scandone, 2020).

We observe anecdotally that several of the partners have found that rigorous evidence is often less available, harder to generate, and harder to leverage in disadvantaged contexts globally. This suggests that prioritizing equity within the program's agenda may require more intentional and specific consideration of it as a **factor in partner selection and monitoring partner activities**. Attempts to overhaul partners' individual approaches to supporting equity would likely be counterproductive; however, monitoring their efforts and ensuring that the overall portfolio of program investments is placing emphasis on remedying inequities that the program deems to be a high priority to address should be feasible.

CHAPTER 4

Implications of Our Findings for the Program and Emerging Next Steps

In the previous chapters, we have described the program's theory of change to date, detailed our findings regarding the program's progress according to key indicators, and presented formative lessons learned with some attendant considerations for the BHP Foundation, based on qualitative data. We have also highlighted examples where we concede that our ability to draw firm conclusions about program progress or impacts is limited, particularly by the evidence available from prespecified indicators. While we are still missing some important data points that will ultimately inform our evaluation—including the program-funded final independent evaluations of each partner's work—this report presents a window into midterm findings at a critical time when the program has begun to plan for the next phase of its investments.

In this chapter, we consider the totality of our findings, drawn from all the indicators and qualitative data. We discuss some potential implications of our combined findings for the program's next iteration. We also provide an update for readers about the ways in which the program leadership, informed by insights and lessons learned from their efforts to date, have already begun to adapt both their investment strategy and their approach to MLE activities. We conclude with brief reflections on the relevance of our midterm findings for the field as a whole and provide an overview of the next steps for the RAND team's evaluation.

Recommendations for the Program Strategy

Informed by our review of the program's strategy and our findings to date, we offer the following midterm recommendations for the program in five key areas.

Recommendation 1. Clarify What Constitutes "Actionable Evidence"

Since it was established, the program's overarching priority has been to invest in partners capable of facilitating a significant *step change* improvement in education. The program has previously defined this as "a behavior change explicitly linked to the use of evidence to inform decision making," characterized by "increased use of evidence across systems (direct or indirect) to improve learning outcomes" (BHP Foundation, 2021). While this focus on evidence capable of improving learning outcomes has been consistent, partners have thus far varied widely in the **standards of evidence** underlying the ideas or practices they are promoting. Some have emphasized interventions proven to work through experimental or quasi-experimental studies that may be replicable at scale, while others have focused on highly contextualized "tacit" insights that are meant to be adapted locally rather than replicated with fidelity. A key question for the program moving forward is whether there is a particular standard of evidence—or perhaps specifying preferred use cases for varying types of evidence depending on the context—that could be clarified as a core feature of the program's evidence-to-practice work. Different types of evidence involve different trade-offs; more subjective and smaller-sample analyses often provide quicker, more-highly contextualized, and timely insights, while

large-scale designed studies can more accurately and objectively assess the feasibility and expected benefits of implementing recommended practices.

Recommendation 2. Set Clearer Goals Around Equity

The program has also consistently sought to address issues of equity, which it has defined broadly as supporting educational outcomes in either low- and middle-income countries or disadvantaged communities within higher-income countries. However, as discussed in Chapter 3, partners vary in the extent and manner in which they address inequities in their work. Given its priority on equity, the program may benefit from being more explicit in the degree to which its funded activities are expected to contribute to advancing equity versus improving educational practice generally. This could involve both a clearer articulation of program goals around equity and more-systematic monitoring of partners' emphasis on equity-related efforts through their funded work. While equity goals are likely to remain multifaceted and inclusive of many high-needs groups, articulating what is in scope, as well as how central equity improvement is to the funded work, would be helpful.

Recommendation 3. Specify Intended Audiences and Anticipated Engagement Models

Program efforts to date have included a significant focus on developing a more robust and user-friendly global evidence base, as well as extensive dissemination and advocacy linked to that evidence. However, the partners have rarely specified the audiences that they expect will be most responsive to their efforts. If the goal is primarily to **curate and disseminate evidence-based practices opportunistically** toward whichever audiences are most inclined to adopt them, this may be sufficient. However, if the program seeks to **facilitate systemwide changes in practice in particular local or national contexts**, funded partners will likely need to better specify the audiences for their work across multiple levels of educational and governance systems, carefully assess demand for the reform agendas being promoted, and tailor their outputs to these audiences. Different partners or potential partners are likely to be more effective in different contexts, and therefore identifying the right match between partners and their intended audiences will be important.

Recommendation 4. Enhance the Rigor of Evaluation Activities and the Selection of Indicators

The independent evaluation mechanisms that the program has supported to date have been effective in gathering stakeholder perspectives and providing formative feedback to partners and the program. These are important contributions, particularly for midterm evaluations. However, partners' independent evaluations have not generally been structured to gather substantial new data or evidence around key components of the work, and their relatively short cycles may limit their capacity to do more than synthesize the previously collected indicators. Meanwhile, while the program initially envisioned key performance indicators as the primary approach to measuring program progress and impact, the indicators used were not clearly aligned to testing specific hypotheses articulated by the program, and the partners have typically not used methods suitable to assessing impacts or efficacy.

A variety of evaluation methodologies could potentially help provide greater clarity about the program's impacts moving forward. To some extent, effects of the program will likely become evident over time as additional prominent **descriptive indicators** of the effects of partners' dissemination and advocacy work are observed and continue to be monitored. In addition, other methods could be leveraged more. Causal **impact analyses**, such as those currently being considered by UNW for its SCE program, could help vali-

date and generate lessons about best practices for a program that is being scaled at an impressive rate. More systematic **assessments of outreach** by multiple partners drawn from surveys and more systematic follow-up of individuals engaged directly via dissemination or capacity-building efforts could gauge the efficacy of partners' activities for informing decisionmaking and empowering thought leaders. Additional targeted **case studies** of partners' flagship efforts that not only include the perspectives of the individuals directly collaborating with partners (as our case studies have done) but that go further to gather data at the ground level where changed practices are expected to occur could also help validate assumed pathways between partners' activities and their anticipated impacts. Finally, analyses that focus on assessing the **core capacities, positioning, or brands** of current or prospective partners could help assess or illustrate their capacity to achieve impact. These analyses may also help gauge the extent to which key intended audiences are likely to be receptive to partner efforts, particularly where the ultimate impacts are expected to be diffuse and not directly measurable.

Recommendation 5. Link Evaluation Activities to a Program Learning Agenda

More-rigorous evaluation activities tend to be more resource- and time-intensive than the program indicator approach used to date. More-rigorous evaluation activities are also difficult to implement on demand and may require additional planning or may need to be embedded into partners' funded work. Therefore, it will be important for the BHP Foundation to **identify those specific assumptions and hypotheses that it views as most essential to test or validate** using more-rigorous methods**.** These may be assumptions central to a particular partner's funded work that also relate to the program's key hypotheses. They might also be proof points that, once established, could be particularly influential or helpful to share with the broader field.

Given that some partners' workstreams have substantially adapted over time, specifying a program learning agenda can also add value by facilitating a more dynamic process for selecting and monitoring indicators of progress. As partner efforts evolve and the need for updated or new indicators arises, these can be selected with an eye toward ensuring sufficient monitoring to test key hypotheses and advance program learning goals.

We defer to the program and partners with respect to which specific areas may be most valuable to explore in greater depth. For example, the BHP Foundation may, in some cases, view a partner's past body of work and evidence as highly credible and determine that additional validation of current funded activities is not needed. At the same time, our examination of the partners' separate theories of change and strands of work suggest several categories of knowledge gaps related to partners' current funded activities, and these may also be relevant to a programwide learning agenda. Key categories of questions are summarized in Table 4.1.

These categories are meant to provide a menu of options for the program to consider as it determines future evaluation priorities at either the partner or program level. The potential methods accompanying these questions could be leveraged as part of the still ongoing final evaluation work with partners in this funding cycle or as part of research and evaluation work linked to future cycles of program investment in similar activities.

Emerging Next Steps for the Program

Some of our midterm recommendations align with steps that program and BHP Foundation leadership are already taking and that are intended to strengthen and clarify the program strategy. This alignment may reflect both the ongoing dialogue between our evaluation team and program leaders and a degree of independent convergence toward some common conclusions. Through interviews with program leaders and review

TABLE 4.1
Key Categories of Knowledge Gaps Related to Partners' Efforts and Potential Methodologies for Addressing Them

Knowledge Gaps	Potential Methodologies
• The extent of demand and take-up for program-funded tools and knowledge resources by end-users	• Case studies; assessments of outreach
• The extent of direct impacts of program-funded educational programs on recipients	• Impact studies
• The utility or reach of findings from program-funded research studies	• Descriptive indicators; case studies
• The extent of knowledge transfer and collaboration within program-facilitated networks	• Case studies; assessments of outreach
• The extent of the impact of advocacy efforts on policymakers, practitioners, or the public	• Indicators; case studies; core capacity assessments
• The extent and durability of investments in local decisionmakers' capacity to use evidence or to scale effective practices	• Indicators; case studies

of documents that they have provided, we have had the opportunity to observe this process and some of the emerging priorities for the program.

First, across all of its funded programs, the BHP Foundation has begun to **refine its approach to MLE**. For the Education Equity program, the new approach to MLE will involve **revisions to its theory of change**, including specification of key hypotheses that it views as most critical to test and examine through its investments and evaluation activities in the next phase of its funded work. This organization-level shift aligns directly to our recommendation to specify a program-wide learning agenda.

Program leaders have also engaged in a **landscape review** (with contributions from the RAND team) to examine key trends in global education relevant to the program strategy. These trends include the ongoing global learning gap that has been exacerbated by the COVID-19 pandemic, trends toward decolonization of development and philanthropy, an increased emphasis on improving education equity and inclusion, an emphasis on whole child development, greater demand for relevant and actionable data and evidence, and more. The program has also engaged in internal reflections about and refinement of the criteria for what is in scope for the program's investments moving forward.

This step-back work on the part of program leaders is still in process; it remains to be seen exactly how the program's approach may shift as a result. However, some key considerations are apparent even at this stage. Program leaders are considering **clarifying their priorities** with respect to investments aimed at enhancing equity and the lenses with which they approach identifying solutions for inequity. In line with this, they have described an increased emphasis on investing in innovative solutions and partners rooted in the global south and engaged in more national (versus global) contexts. They have also discussed the possibility of reorienting program efforts toward evidence that provides more immediately applicable implementation guidance for decisionmakers versus more academically oriented research. Some other aspects of their strategy appear likely to remain the same, including the focus on investments in long-term systemic improvements and on testing new and potentially transformative innovations.

CHAPTER 5

Broader Implications of This Work and Next Steps for the Evaluation

Consistent with the BHP Foundation's goals in commissioning an independent and transparent evaluation, one purpose of this report is to inform the program's continuous improvement efforts. In addition, the evaluation and this report are also intended to help inform the field as a whole about valuable lessons learned from its funded work. Both some of the **project-specific insights** highlighted in our case studies and some **overarching findings** about the program's progress may be relevant to other actors engaged in similar evidence-to-practice work globally.

First, partners' funded activities in the areas of **curation and dissemination** suggest key considerations for other organizations seeking to promote the use of research and extant knowledge to inform educational practices in diverse contexts. These considerations include the importance of contextualizing evidence to increase transferability, solicitation of local perspectives to ensure relevance and applicability of what is curated, and, in some cases, finding a middle ground between recommending what has been proven to work in the past in one context and recommending related efforts that appear to be working in another. The partners' wide-ranging efforts also provide useful examples of what works when promoting evidence-based reforms to various audiences spanning policymakers, practitioners, and business leaders. These examples are discussed in greater detail in our case study investigations, as well as in publications shared by the partners themselves (Wolfe, Wang, and Master, 2023a, -b, -c; Master et al., 2023; GBC-Ed, 2022; Robinson, Wyss, and Hannahan, 2021). Overall, while the global evidence base for educational improvement is large and growing, the ability to leverage it to promote improved practices locally may depend crucially on efforts to **better apply that knowledge to new circumstances** and to communicate more effectively with diverse audiences.

Second, the program's efforts highlight the **importance of the networks of collaboration** that partners had previously established and/or developed to facilitate their work. As described in Chapter 3 and detailed in our case study investigations, partners heavily rely on collaborations with and between a range of brokers, clients, and practitioners to maximize their impact. Scaling up such educational interventions as the SCE program, gathering information about promising innovations, monitoring whether tools and resources are relevant to end-users, and facilitating ongoing adoption and adaptation of recommended practices are all key aspects of this work that depend on partners' wide-ranging networks of affiliation and collaboration. While it is not yet clear which types of partner networks will prove to be the most useful for facilitating educational improvements at scale, their collective work to date points to networks of purposeful collaboration as being critical to ensuring the reach, relevance, and longevity of their contributions.

Finally, we have also noted the very **long lead time** associated with investments in evidence-to-practice work of this type, a finding that is not unexpected but is still an important consideration for other organizations engaged in this space. In broad strokes, we have observed that the partners have been initially effective in activities that they could directly control. These include developing global public goods, such as evidence repositories, toolkits for policymakers or decisionmakers, and direct or partner-mediated provision of educational services at scale. While these efforts have significant potential for impact, our program evaluation

highlights how leveraging them to help facilitate systemic reforms in educational practices will take time and **often depends on a series of contingent next steps**. Clearer articulation of envisioned long-term pathways to impact and more-rigorous monitoring of near-term proof points to pressure test those hypotheses will be important if actors are to identify which evidence-to-practice innovations are capable of contributing to step changes in educational outcomes.

Next Steps for the RAND Team's Evaluation

Our next and **final evaluation report** about the Education Equity program, to be published in fall 2024, will include **summative findings** about the impacts and insights gained from the program's first phase of investments. The report will be informed by assessments of partners' work and progress against indicators through the end of their initially funded work, by program-commissioned independent final evaluations of their progress, and by the full suite of interviews and case studies that we are conducting to gain insight into key strands of partners' efforts. We will also have the opportunity to describe in greater detail **how the program strategy has evolved** over time and how this translates into its priorities for investments in the next stage of its work in the area of global education. In addition to describing key trends and the observed impacts of the evidence-to-practice work funded by the program, we also anticipate that the program's arc will provide a valuable reference point for the field. Other actors in this space may benefit from **learning both what it has accomplished and how it has adapted** in response to lessons learned and to the shifting landscape of global educational challenges since its inception.

APPENDIX

Key Indicators Tables

This appendix provides data for the key indicators that we identified from the Education Equity program partners linked to the activities and interim outcomes of the program-level theory of change. For both the activities tables (A.1–A.5) and the interim outcomes tables (A.6–A.9), we list the partner from which the indicator was drawn, the indicator, the partner-provided prespecified target (if available), the status of progress for the indicator as documented in the partner's reporting, and the date through which the data are current (typically June or September 2022). For interim outcome indicators, we also include our characterization of which activities in the theory of change most directly lead to the interim outcome being measured.

Often, we view the same indicator as a relevant signal of progress for multiple activities or interim outcomes, and, in those cases, the tables for each activity or interim outcome list the other activities or interim outcomes that the indicator measures. We do, however, restrict each indicator to one phase of the theory of change, either measuring project activities or interim outcomes resulting from those activities. This is despite some fluidity in where to draw the line.

The indicator framework is designed to shed light on the partners' collective progress in conducting activities and achieving interim outcomes aligned with the program theory of change. Because of the varied nature of the partners' work and the pathways they are pursuing to achieve impact, some partners have more indicators for certain activities or interim outcomes than others, and some have none in some areas. This is to be expected and reflects the diverse range of projects that the BHP Foundation funded through the program. The indicator framework is not intended to replace the partners' individual project evaluations, and it would not be appropriate to evaluate individual partners solely based on the key indicators included here.

TABLE A.1

Key Activity Indicators: Curate Evidence About Improving Education

Partner	Indicator	Also Indicator for Activities	Target	Status	Last Updated
CUE	Number of knowledge products (publications, tools, podcasts, etc.) developed by the Innovations team	Disseminate evidence	N/A	27 knowledge products (Y1–Y2); 34 knowledge products (Y3–Y4)	June 2022
CUE	Number of knowledge products (publications, tools, podcasts, etc.) developed by the Analyze Data team	Disseminate evidence	N/A	16 knowledge products (Y1–Y2); 21 knowledge products (Y3–Y4)	June 2022
CUE	Number of knowledge products (publications, tools, podcasts, etc.) developed by the Scaling team	Disseminate evidence	N/A	27 knowledge products (Y1–Y2); 23 knowledge products (Y3–Y4)	June 2022
EEF	Metrics on Toolkit development: number of Toolkit studies coded; number of reviews for Toolkit strands completed	Disseminate evidence	No explicit coding target; 34 reviews by March 2021	More than 2,700 studies coded across target of 34 strands prior to release of Toolkit 2.0 in September 2021; since then, EEF has begun automatic searching for additional studies for 13 strands to make Toolkit a "living review"; updating and integrating Early Years Toolkit into main Toolkit also underway	September 2022
EEF	Number of project fellows and number of country or regional jurisdictions with fellows	Develop capacity	N/A	11 fellows to date across seven jurisdictions	September 2022
GBC-Ed	Metrics on breadth of global education landscape analyses (e.g., number of studies, major actors, regulatory frameworks, guiding principles, strategic initiatives assessed and analyzed) or the number of stakeholders consulted	N/A	Interviews with 15–20 business leaders and 15 noncorporate entities; no other explicit targets	Interviewed 46 business leaders and 45 noncorporate entity leaders as part of its scoping, listening, and designing work, exceeding target; evidence curation work ongoing to support tool and resource development	September 2022
TFALL	Metrics on volume of Community Impact Lab learning experiences: number of learning experiences; number of participants in learning experiences	Disseminate evidence; develop capacity	Three of each type of experience annually by 2020–2021: learning loops, virtual communities of practice, learning teams, expert practitioner roundtables, and virtual workshops	Learning experiences ongoing, including 30 with 900 participants in 2020–2021, eight with 1,300 participants through midyear 2021–2022, and additional roundtables and fellowship sessions over April–June 2022; evolution of types of learning experiences being offered makes it difficult to assess relative to its target	June 2022

Table A.1—Continued

Partner	Indicator	Also Indicator for Activities	Target	Status	Last Updated
TFALL	Number of resources related to community impact housed in online resource bank (e.g., digital papers, case studies, snapshots, and artifacts)	Disseminate evidence	At least 50 resources curated and published annually beginning in 2019–2020	Added more than 50 resources related to community impact to its online resource bank per year through 2021–2022, meeting its target; continues to compile additional resources with plans to publish them	June 2022
TFALL	Number of resources and tools highlighting individual partner practices in alumni leadership development curated and posted on to Partner Learning Portal	Disseminate evidence	Three collective leadership resource bundles, five self-guided learning materials, and two new exemplars in each of the five core areas of leadership development added annually beginning in 2019–2020	Resources highlighting partner practices in leadership development being developed, including "contextualized resources" being created by early adopters of the "Teaching as Collective Leadership" model; numbers and types of resources vary from initial vision, making it difficult to assess relative to its target	June 2022
UNW	Number of learners accessing high-quality SCE content	Test innovations	67,000 women participating in the SCE program via face-to-face or online learning	101,179 women have accessed SCE content, exceeding its target for the pilot period	September 2022
UNW	Number of training modules and curricula developed and informed by SCE country needs assessments	Test innovations	21 modules	529 modules developed, exceeding its target for the pilot period	September 2022

NOTE: N/A in "Also Indicator for Activities" column = no other activities for which an indicator; N/A in "Target" column = no prespecified target; Y = year.

TABLE A.2

Key Activity Indicators: Disseminate Evidence-Based Strategies

Partner	Indicator	Also Indicator for Activities	Target	Status	Last Updated
CUE	Number of knowledge products (publications, tools, podcasts, etc.) developed by the Innovations team	Curate evidence	N/A	27 knowledge products (Y1–Y2); 34 knowledge products (Y3–Y4)	June 2022
CUE	Number of knowledge products (publications, tools, podcasts, etc.) developed by the Analyze Data team	Curate evidence	N/A	16 knowledge products (Y1–Y2); 21 knowledge products (Y3–Y4)	June 2022
CUE	Number of knowledge products (publications, tools, podcasts, etc.) developed by the Scaling team	Curate evidence	N/A	27 knowledge products (Y1–Y2); 23 knowledge products (Y3–Y4)	June 2022
CUE	Number of CUE-initiated consultations and events related to the work of the Innovations team	Develop capacity; advocacy	N/A	275 consultations and 19 events (Y1–Y2); 106 consultations and 31 events (Y3–Y4)	June 2022
CUE	Number of CUE-initiated consultations and events related to the work of the Analyze Data team	Develop capacity; advocacy	N/A	84 consultations and six events (Y1–Y2); 45 consultations and 17 events (Y3–Y4)	June 2022
CUE	Number of CUE-initiated consultations and events related to the work of the Scaling team	Develop capacity; advocacy	N/A	154 consultations and 24 events (Y1–Y2); 96 consultations and 23 events (Y3–Y4)	June 2022
EEF	Metrics on Toolkit development: number of Toolkit studies coded; number of reviews for Toolkit strands completed	Curate evidence	No explicit coding target; 34 reviews by March 2021	More than 2,700 studies coded across target of 34 strands prior to release of Toolkit 2.0 in September 2021; since then, EEF has begun automatic searching for additional studies for 13 strands to make Toolkit a "living review"; updating and integrating Early Years Toolkit into main toolkit also underway	September 2022

Table A.2—Continued

Partner	Indicator	Also Indicator for Activities	Target	Status	Last Updated
EEF	Number of evidence hubs established	Develop capacity	One in each partner jurisdiction	One partner jurisdiction with expressed interest in evidence hub model and 0 with established hubs, falling short of target; five developing alternative contextual evidence mobilization approaches; pilot evidence hub networks in two jurisdictions that are not core partners for BHP Foundation work	September 2022
GBC-Ed	Metrics on tool development: number of prototypes, number finalized	Develop capacity	Three to five prototypes, at least one finalized tool	Has begun to roll out a set of tools and resources for businesses, such as services at no cost to "Coalition" members and several paid tiers of services; nature of a *tool* has evolved such that it is difficult to quantify the number prototyped or finalized and assess relative to target	September 2022
TFALL	Metrics on volume of Community Impact Lab learning experiences: number of learning experiences; number of participants in learning experiences	Curate evidence; develop capacity	Three of each type of experience annually by 2020–2021: learning loops, virtual communities of practice, learning teams, expert practitioner roundtables, and virtual workshops	Learning experiences ongoing, including 30 with 900 participants in 2020–2021, eight with 1,300 participants through midyear 2021–2022, and additional roundtables and fellowship sessions over April–June 2022; evolution of types of learning experiences being offered makes assessment relative to its target difficult	June 2022
TFALL	Number of resources related to community impact housed in online resource bank (e.g., digital papers, case studies, snapshots, and artifacts)	Curate evidence	At least 50 resources curated and published annually beginning in 2019–2020	Added more than 50 resources related to community impact to its online resource bank per year through 2021–2022, meeting its target; continues to compile additional resources with plans to publish them	June 2022

Table A.2—Continued

Partner	Indicator	Also Indicator for Activities	Target	Status	Last Updated
TFALL	Number of resources and tools highlighting individual partner practices in alumni leadership development curated and posted onto Partner Learning Portal	Curate evidence	Three collective leadership resource bundles, five self-guided learning materials, and two new partner exemplars in each of the five core areas of leadership development added annually beginning in 2019–2020	Resources highlighting partner practices in leadership development being developed, including "contextualized resources" being created by early adopters of the "Teaching as Collective Leadership" model; numbers and types of resources vary from initial vision, making assessment relative to its target difficult	June 2022
TFALL	Metrics on activities designed to advance TFALL partner learning: number of partner support initiatives undertaken by TFALL specialists; number of network conferences and group learning experiences offered	Develop capacity	Four virtual learning communities for ten or more partner staff, ten in-person support initiatives to strengthen partner staff capacity, and 20 remote support initiatives annually beginning in 2019–2020	Activities to advance partner learning are ongoing, including via development of an alumni strategy guide and 25–30 projects related to participant training, student impact, or alumni strategy in at least five regions; difficult to assess relative to its target due to project evolution	June 2022
UNW	Number of policy dialogues on the importance of women's right to education and vocational learning undertaken with UNW support, country and global level	Advocacy	26 country-level dialogues; two global dialogues	96 policy dialogues in individual countries during the pilot period, exceeding its target; no global policy dialogue to date	September 2022

NOTE: N/A in "Also Indicator for Activities" column = no other activities for which an indicator; N/A in "Target" column = no prespecified target; Y = year.

TABLE A.3

Key Activity Indicators: Develop Capacity to Leverage Evidence

Partner	Indicator	Also Indicator for Activities	Target	Status	Last Updated
CUE	Number of CUE-initiated consultations and events related to the work of the Innovations team	Disseminate evidence; advocacy	N/A	275 consultations and 19 events (Y1–Y2); 106 consultations and 31 events (Y3–Y4)	June 2022
CUE	Number of CUE-initiated consultations and events related to the work of the Analyze Data team	Disseminate evidence; advocacy	N/A	84 consultations and six events (Y1–Y2); 45 consultations and 17 events (Y3–Y4)	June 2022
CUE	Number of CUE-initiated consultations and events related to the work of the Scaling team	Disseminate evidence; advocacy	N/A	154 consultations and 24 events (Y1–Y2); 96 consultations and 23 events (Y3–Y4)	June 2022
EEF	Number of project fellows and number of country/regional jurisdictions with fellows	Curate evidence	N/A	11 fellows to date across seven jurisdictions	September 2022
EEF	Number of evidence hubs established	Disseminate evidence	One in each partner jurisdiction	One partner jurisdiction with expressed interest in evidence hub model and 0 with established hubs, falling short of target; five developing alternative contextual evidence mobilization approaches; pilot evidence hub networks in two jurisdictions that are not core partners for BHP Foundation work	September 2022
GBC-Ed	Metrics on tool development: number of prototypes, number finalized	Disseminate evidence	Three to five prototypes, at least one finalized tool	Has begun to roll out a set of tools and resources for businesses, such as services at no cost to "Coalition" members and several paid tiers of services; nature of a *tool* has evolved such that it is difficult to quantify the number prototyped or finalized and assess relative to target	September 2022
TFA	Metrics on volume of Community Impact Lab learning experiences: number of learning experiences; number of participants in learning experiences	Curate evidence; disseminate evidence	Three of each type of experience annually by 2020–2021: learning loops, virtual communities of practice, learning teams, expert practitioner roundtables, and virtual workshops	Learning experiences ongoing, including 30 with 900 participants in 2020–2021, eight with 1,300 participants through midyear 2021–2022, and additional roundtables and fellowship sessions over April–June 2022; evolution of types of learning experiences being offered makes it difficult to assess relative to its target	June 2022

Table A.3—Continued

Partner	Indicator	Also Indicator for Activities	Target	Status	Last Updated
TFA	Metrics on activities designed to advance TFALL partner learning: number of partner support initiatives undertaken by TFALL specialists; number of network conferences and group learning experiences offered	Disseminate evidence	Four virtual learning communities for ten or more partner staff, ten in-person support initiatives to strengthen partner staff capacity, and 20 remote support initiatives annually beginning in 2019–2020	Activities to advance partner learning are ongoing, including via development of an alumni strategy guide and 25–30 projects related to participant training, student impact, or alumni strategy in at least five regions; difficult to assess relative to its target due to project evolution	June 2022
TFA	Number of communities participating in community-level progress matrix and that have established local indicators for progress	N/A	50 communities experimenting with approaches to measure community-level progress by 2020–2021, with indicators for measuring progress identified in a similar number of communities by 2021–2022	Pivoted its monitoring and evaluation approach away from community-level progress matrix approach, thereby falling short of its target to have 50 communities develop progress indicators as part of that matrix by 2021–2022	June 2022
UNW	Number of physical spaces or Women's Empowerment Hubs or community spaces providing access to SCE e-learning platform and other SCE education learning opportunities established with UNW support	Test innovations	45 physical spaces	83 physical spaces or Women's Empowerment hubs where SCE programs are available established, exceeding its target	September 2022

NOTE: N/A in "Also Indicator for Activities" column = no other activities for which an indicator; N/A in "Target" column = no prespecified target; Y = year.

TABLE A.4

Key Activity Indicators: Advocate to Develop Enabling Environments

Partner	Indicator	Also Indicator for Activities	Target	Status	Last Updated
CUE	Number of CUE-initiated consultations and events related to the work of the Innovations team	Disseminate evidence; develop capacity	N/A	275 consultations and 19 events (Y1–Y2); 106 consultations and 31 events (Y3–Y4)	June 2022
CUE	Number of CUE-initiated consultations and events related to the work of the Analyze Data team	Disseminate evidence; develop capacity	N/A	84 consultations and six events (Y1–Y2); 45 consultations and 17 events (Y3–Y4)	June 2022
CUE	Number of CUE-initiated consultations and events related to the work of the Scaling team	Disseminate evidence; develop capacity	N/A	154 consultations and 24 events (Y1–2); 96 consultations and 23 events (Y3–Y4)	June 2022
EEF	Number of engagements with policymakers	N/A	N/A	Reports that all partners have "a strong working relationship" with at least one jurisdiction; lists ten countries as "recent successes" in building robust relationships with policymakers; multiple recent policymaker engagements in Australia	September 2022
UNW	Number of policy dialogues on the importance of women's right to education and vocational learning undertaken with UNW support, country and global level	Disseminate evidence	26 country-level dialogues; two global dialogues	96 policy dialogues in individual countries during the pilot period, exceeding its target, no global policy dialogue to date	September 2022
UNW	Number of awareness and advocacy campaigns on the importance of young women's right to education and vocational learning undertaken with UNW support, country and global level	N/A	13 country-level campaigns, three global campaigns	66 country-level awareness and advocacy campaigns, exceeding its target; no global awareness and advocacy campaign to date	September 2022

NOTE: N/A in "Also Indicator for Activities" column = no other activities for which an indicator; N/A in "Target" column = no prespecified target; Y = year.

TABLE A.5

Key Activity Indicators: Test Potentially Transformative Innovations

Partner	Indicator	Also Indicator for Activities	Target	Status	Last Updated
EEF	Metrics on volume of trials: number of applications to conduct trials; number of trials conducted	N/A	Two trials in each partner jurisdiction	Ten funded trials in five jurisdictions (one not a core partner jurisdiction), but not yet at target of two per partner jurisdiction; trials substantially delayed due to COVID-19 and difficulties identifying matching funds	September 2022
UNW	Number of learners accessing high-quality SCE content	Curate evidence	67,000 women participating in the Second Chance Education program via face-to-face or online learning	101,179 women have accessed SCE content, exceeding its target for the pilot period	September 2022
UNW	Number of training modules or curricula developed and informed by SCE country needs assessments	Curate evidence	21 modules	529 modules developed, exceeding its target for the pilot period	September 2022
UNW	Number of physical spaces or Women's Empowerment Hubs or community spaces providing access to SCE e-learning platform and other SCE learning opportunities established with UNW support	Develop capacity	45 physical spaces	83 physical spaces or Women's Empowerment hubs where SCE programs are available established, exceeding its target	September 2022

NOTE: N/A in "Also Indicator for Activities" column = no other activities for which an indicator; N/A in "Target" column = no prespecified target.

TABLE A.6

Key Interim Outcome Indicators: Increased Availability of Data and Evidence

Partner	Indicator	Also Indicator for Interim Outcomes	Activities Leading to Interim Outcome	Target	Status	Last Updated
CUE	Number of externally initiated engagements related to the work of the Innovations team, such as requests (participated in and declined) from the media, briefings, panels, etc.	Empowered leaders and actors	Curate evidence; disseminate evidence; develop capacity; advocacy	N/A	93 engagements (Y1–Y2); 99 engagements (Y3–Y4)	June 2022
CUE	Number of externally initiated engagements related to the work of the Analyze Data team, such as requests (participated in and declined) from the media, briefings, panels, etc.	Empowered leaders and actors	Curate evidence; disseminate evidence; develop capacity; advocacy	N/A	78 engagements (Y1–Y2); 94 engagements (Y3–Y4)	June 2022

Table A.6—Continued

Partner	Indicator	Also Indicator for Interim Outcomes	Activities Leading to Interim Outcome	Target	Status	Last Updated
CUE	Number of externally initiated engagements related to the work of the Scaling team, such as requests (participated in and declined) from the media, briefings, panels, etc.	Empowered leaders and actors	Curate evidence; disseminate evidence; develop capacity; advocacy	N/A	34 engagements (Y1–Y2); 20 engagements (Y3–Y4)	June 2022
EEF	Rate of use of Toolkit to make resource allocation or school improvement decisions, as measured by proportion of school-based and system leadership staff reporting decisionmaking based on Toolkit evidence	Informed decisionmaking	Curate evidence; disseminate evidence	N/A	More than 850,000 page views from Toolkit 2.0 launch through March 2022 (metric not provided in more-recent reporting); unclear whether there has been improvement in the rate at which school-based and system leadership staff are using Toolkit evidence to inform decisionmaking	September 2022
EEF	Number of publications of trial results	N/A	Test innovations	N/A	No published results from trials yet because of delays in conducting them	September 2022
TFALL	Reach of stories and reports prepared for dissemination in public channels as measured by number of media mentions and number of attendees at public events	N/A	Curate evidence; disseminate evidence	N/A	Reports “several high profile mentions in external media and events” but has not yet developed standardized metrics for tracking the reach of stories and reports prepared for public dissemination	June 2022
UNW	Number and percentage increase in employed or entrepreneur women among program graduates	Empowered leaders and actors	Curate evidence; develop capacity; test innovations	17,500 women	32,884 employed or entrepreneur women among program graduates, exceeding target	September 2022

NOTE: N/A in “Also Indicator for Interim Outcomes” column = no other interim outcomes for which an indicator; N/A in “Target” column = no prespecified target; Y = year.

TABLE A.7

Key Interim Outcome Indicators: More-Informed Decisionmaking

Partner	Indicator	Also Indicator for Interim Outcomes	Activities Leading to Interim Outcome	Target	Status	Last Updated
EEF	Rate of use of Toolkit to make resource allocation or school improvement decisions, as measured by proportion of school-based and system leadership staff reporting decisionmaking based on Toolkit evidence	Availability of data and evidence	Curate evidence; disseminate evidence	N/A	More than 850,000 page views from Toolkit 2.0 launch through March 2022 (metric not provided in more recent reporting); unclear whether there has been improvement in the rate at which school-based and system leadership staff are using Toolkit evidence to inform decisionmaking	September 2022
EEF	Metrics on the translation of engagement with policymakers into policy	Empowered leaders and actors	Advocacy	N/A	Toolkit was used to guide €5.8 billion in postpandemic school spending in the Netherlands; ongoing partner engagement with four African governments to identify schools where evidence-based practices could be incorporated into teacher training	September 2022
GBC-Ed	Metrics on use of tools such as the number of business users, number of workshops and training sessions for ESG professionals, or the number of professionals trained	Empowered leaders and actors	Curate evidence; disseminate evidence; develop capacity	Ten to 15 companies and ten noncorporate entities using tools developed	Engaged 52 business leaders from 43 companies in its no-cost community, exceeding target of ten to 15 businesses using its tools, though intensity of engagement is unclear	September 2022

Table A.7—Continued

Partner	Indicator	Also Indicator for Interim Outcomes	Activities Leading to Interim Outcome	Target	Status	Last Updated
GBC-Ed	Metrics related to dissemination and implementation of guidance material for noncorporate stakeholders in education	Empowered leaders and actors	Disseminate evidence; advocacy	Surveys/ interviews of private sector engagement team/staff of three major global education funds, ten UN agencies and international organizations, and ten donor and developing country partners	Guidance material for noncorporate stakeholders has not yet been produced and the organization has not yet met its target for engaging this sector	September 2022
TFA	Network partner staff and global stakeholders report that their knowledge increased after engagements with the Community Impact Lab as measured by responses to surveys	Empowered leaders and actors	Curate evidence; disseminate evidence; develop capacity	N/A; seeking improvement over time	Surveys show that 90 percent of attendees report the experiences as relevant and can apply knowledge or skills from the experience; the most recent figures increased slightly from the prior year	June 2022
UNW	Number of new, revised, or in active review or discussion legislative frameworks, policies, or revised budgetary allocations that promote SCE and vocational training opportunities for women developed and/or being implemented in SCE countries	Empowered leaders and actors	Disseminate evidence; advocacy	Nine frameworks or policies (general); six revised budgetary allocations	20 revised or in active review legislative frameworks or policies related to SCE, exceeding target; two revised budgetary frameworks, not yet meeting target	September 2022

NOTE: N/A in "Also Indicator for Interim Outcomes" column = no other interim outcomes for which an indicator; N/A in "Target" column = no prespecified target.

TABLE A.8

Key Interim Outcome Indicators: Networks of Purposeful Collaboration

Partner	Indicator	Also Indicator for Interim Outcomes	Activities Leading to Interim Outcome	Target	Status	Last Updated
EEF	Number of annual engagements between evidence hubs and supported schools (both formal meetings and otherwise)	Empowered leaders and actors	Disseminate evidence; develop capacity	50 formal engagements per evidence hub annually	Has not executed its plan to develop evidence hub networks in each partner jurisdiction and is therefore falling short of its target of evidence hubs having 50 formal engagements per year with supported schools in its network	September 2022
EEF	Number of organizations supported to take on the role of evidence brokers	N/A	Develop capacity	N/A	Seeking to consolidate and develop branding for its existing network of partners as the "Evidence for Education Network"; internal SharePoint site created and external website being developed; cultivating relationships with several potential partners ranging from European countries to India and the Eastern Caribbean states	September 2022
TFALL	Reach of TFALL alumni network as measured by total number of alumni, percentage working in education, and percentage working in roles that expand opportunities for children in high-need communities	Empowered leaders and actors	Develop capacity	N/A; seeking improvement over time	Has grown its network to more than 88,000 program alumni and reports, based on survey data, that 70 percent are working or studying in fields related to its mission	June 2022
TFALL	Engagement of TFALL alumni network as measured by number and percentage of alumni responding to survey, percentage increase in mindsets associated with effective alumni leadership, and percentage increase in Net Promoter Score	Empowered leaders and actors	Disseminate evidence; develop capacity	N/A; seeking improvement over time	Percentage of survey respondents reporting agreement with "mindset" questions has been around 80 percent for the past few years, while surveyed alumni and alumni response rates have varied	June 2022

Table A.8—Continued

Partner	Indicator	Also Indicator for Interim Outcomes	Activities Leading to Interim Outcome	Target	Status	Last Updated
UNW	Number of partnerships established identifying community spaces and empowerment hubs with local governments, communities, and private sector partners to create safe learning spaces	N/A	Develop capacity; test innovations	70 partnerships	156 partnerships, exceeding target	September 2022
UNW	Number of peer networks for women established with UNW support, country and global level	N/A	Develop capacity; test innovations	21 country-level networks, one global network	178 country-level peer networks of women, exceeding target, but no global peer network established yet	September 2022

NOTE: N/A in "Also Indicator for Interim Outcomes" column = no other interim outcomes for which an indicator; N/A in "Target" column = no prespecified target.

TABLE A.9

Key Interim Outcome Indicators: Empowered Thought Leaders and Local Actors

Partner	Indicator	Also Indicator for Interim Outcomes	Activities Leading to Interim Outcome	Target	Status	Last Updated
CUE	Number of externally initiated engagements related to the work of the Innovations team, such as requests (participated in and declined) from the media, briefings, panels, etc.	Availability of data and evidence	Curate evidence; disseminate evidence; develop capacity; advocacy	N/A	93 engagements (Y1–Y2); 99 engagements (Y3–Y4)	June 2022
CUE	Number of externally initiated engagements related to the work of the Analyze Data team, such as requests (participated in and declined) from the media, briefings, panels, etc.	Availability of data and evidence	Curate evidence; disseminate evidence; develop capacity; advocacy	N/A	78 engagements (Y1–Y2); 94 engagements (Y3–Y4)	June 2022

Table A.9—Continued

Partner	Indicator	Also Indicator for Interim Outcomes	Activities Leading to Interim Outcome	Target	Status	Last Updated
CUE	Number of externally initiated engagements related to the work of the Scaling team, such as requests (participated in and declined) from the media, briefings, panels, etc.	Availability of data and evidence	Curate evidence; disseminate evidence; develop capacity; advocacy	N/A	34 engagements (Y1–Y2); 20 engagements (Y3–Y4)	June 2022
EEF	Number of annual engagements between evidence hubs and supported schools (both formal meetings and otherwise)	Networks established	Disseminate evidence; develop capacity	50 formal engagements per evidence hub annually	Has not executed its plan to develop evidence hub networks in each partner jurisdiction and is therefore falling short of its target of evidence hubs having 50 formal engagements per year with supported schools in its network	September 2022
EEF	Metrics on the translation of engagement with policymakers into policy	Informed decisionmaking	Advocacy	N/A	Toolkit was used to guide €5.8 billion in postpandemic school spending in the Netherlands; ongoing partner engagement with four African governments to identify schools where evidence-based practices could be incorporated into teacher training	September 2022
GBC-Ed	Metrics on use of tools, such as the number of business users, number of workshops and training sessions for ESG professionals, or the number of professionals trained	Informed decisionmaking	Curate evidence; disseminate evidence; develop capacity	Ten to 15 companies and ten noncorporate entities using tools developed	Engaged 52 business leaders from 43 companies in its no-cost community, exceeding target of ten to 15 businesses using its tools, though intensity of engagement is unclear	September 2022

Table A.9—Continued

Partner	Indicator	Also Indicator for Interim Outcomes	Activities Leading to Interim Outcome	Target	Status	Last Updated
GBC-Ed	Metrics related to dissemination and implementation of guidance material for noncorporate stakeholders in education	Informed decisionmaking	Disseminate evidence; advocacy	Surveys and interviews of private-sector engagement team or staff of three major global education funds, ten UN agencies and international organizations, and ten donor and developing-country partners	Guidance material for noncorporate stakeholders has not yet been produced, and the organization has not yet met its target for engaging this sector	September 2022
TFALL	Network partner staff and global stakeholders report that their knowledge increased after engagements with the Community Impact Lab as measured by responses to surveys	Informed decisionmaking	Curate evidence; disseminate evidence; develop capacity	N/A; seeking improvement over time	Surveys show that 90 percent of attendees report the experiences as relevant and can apply knowledge or skills from the experience; the most recent figures increased slightly from the prior year	June 2022
TFALL	Reach of TFALL alumni network as measured by total number of alumni, percentage working in education, and percentage working in roles that expand opportunities for children in high-need communities	Networks established	Develop capacity	N/A; seeking improvement over time	Has grown its network to more than 88,000 program alumni and reports, based on survey data, that 70 percent are working or studying in fields related to its mission	June 2022
TFALL	Engagement of TFALL alumni network as measured by number and percentage of alumni responding to survey, percentage increase in mindsets associated with effective alumni leadership, and percentage increase in Net Promoter Score	Networks established	Disseminate evidence; develop capacity	N/A; seeking improvement over time	Percentage of survey respondents reporting agreement with "mindset" questions has been around 80 percent for the past few years, while surveyed alumni and alumni response rates have varied	June 2022

Table A.9—Continued

Partner	Indicator	Also Indicator for Interim Outcomes	Activities Leading to Interim Outcome	Target	Status	Last Updated
UNW	Number and percentage increase in employed or entrepreneur women among program graduates	Availability of data and evidence	Curate evidence; develop capacity; test innovations	17,500 women	32,884 employed or entrepreneur women among program graduates, exceeding target	September 2022
UNW	Number of new, revised, or in active review or discussion legislative frameworks, policies, or revised budgetary allocations that promote SCE and vocational training opportunities for women developed and/or being implemented in SCE countries	Informed decisionmaking	Disseminate evidence; advocacy	Nine frameworks or policies (general); six revised budgetary allocations	20 revised or in active review legislative frameworks or policies related to SCE, exceeding target; two revised budgetary frameworks, not yet meeting target	September 2022

NOTE: N/A in "Also Indicator for Interim Outcomes" column = no other interim outcomes for which an indicator; N/A in "Target" column = no prespecified target; Y = year.

Abbreviations

COVID-19	coronavirus disease 2019
CUE	Center for Universal Education at the Brookings Institution
EEF	Education Endowment Foundation
ESG	environmental, social, and governance
FEEN	Family Engagement in Education Network
GBC-Ed	Global Business Coalition for Education
MLE	measurement, learning, and evaluation
SCE	Second Chance Education
TFALL	Teach for All
UNESCO	United Nations Educational, Scientific, and Cultural Organization
UNW	UN Women

References

Bamberger, Michael, and Eleanor Hewitt, *Monitoring and Evaluating Urban Development Programs: A Handbook for Program Managers and Researchers*, World Bank, 1986.

Bernard, H. Russell, Amber Wutich, and Gery W. Ryan, *Analyzing Qualitative Data: Systematic Approaches*, SAGE Publications, 2016.

BHP Foundation, "Education Equity," program homepage, undated. As of April 10, 2023: https://www.bhp-foundation.org/en/programs-projects/education-equity/

BHP Foundation, "Response to the BHP Foundation Education Equity Global Signature Program Baseline Evaluation Report—*Transforming Global Education Through Evidence*," unpublished memorandum, March 2021.

Braun, V., and V. Clarke, "Using Thematic Analysis in Psychology," *Qualitative Research in Psychology*, Vol. 3, No. 2, 2006.

Byrne, D., "A Worked Example of Braun and Clarke's Approach to Reflexive Thematic Analysis," *Quality & Quantity*, Vol. 56, No. 3, 2022.

Denzin, Norman K., and Yvonna S. Lincoln, *Collecting and Interpreting Qualitative Materials*, SAGE Publications, 2003.

GBC-Ed—*See* Global Business Coalition for Education.

Global Business Coalition for Education, *Unlocking Potential and Performance: Recognizing Education's Position at the Core of ESG*, 2022. As of March 9, 2023: https://gbc-education.org/resources/environmental-social-governance/?gclid=EAIaIQobChMInoytr6fP_QIV7smUCR3gCQk2EAAYASAAEgKeN_D_BwE

Haims, Marla C., Melinda Moore, Harold D. Green, Jr., and Cynthia C. Clapp-Wincek, *Developing a Prototype Handbook for Monitoring and Evaluating Department of Defense Humanitarian Assistance Projects*, RAND Corporation, TR-784-OSD, 2011. As of April 7, 2023: https://www.rand.org/pubs/technical_reports/TR784.html

Husain, Fatima, and Berenice Scandone, *Building a Global Evidence Ecosystem for Teaching Evaluation: Phase 1 Report*, unpublished project evaluation, July 2020.

Master, Benjamin K., Shelly Culbertson, Brian Phillips, Elaine Lin Wang, Harold D. Green, Joe Francombe, Hamish Evans, and Susan Guthrie, *Transforming Global Education Through Evidence: An Evaluation System for the BHP Foundation's Education Equity Global Signature Program,* RAND Corporation, RR-A239-1, 2021. As of April 7, 2023: https://www.rand.org/pubs/research_reports/RRA239-1.html

Master, Benjamin K., Harold D. Green, Brian Phillips, and Elaine Lin Wang, *Exploring How Teach For All's Networks Connect Local Educational Organizations: A Case Study of Alumni Professional Connections in Peru*, RAND Corporation, RR-A259-7, 2023.

McDavid, James C., and Laura R. L. Hawthorn, *Program Evaluation and Performance Measurement: An Introduction to Practice*, SAGE Publications, 2006.

Robinson, Jenny Perlman, Molly Curtiss Wyss, and Patrick Hannahan, *Putting Scaling Principles into Practice: Resources to Expand and Sustain Impact in Education*, Brookings Institution, July 7, 2021. As of March 9, 2023: https://www.brookings.edu/research/putting-scaling-principles-into-practice-resources-to-expand-and-sustain-impact-in-education/

Roser, Max, and Esteban Ortiz-Ospina, "Global Education," webpage, *Our World In Data*, 2016. As of June 3, 2022: https://ourworldindata.org/global-education

Ryan, Gery W., and H. Russell Bernard, "Techniques to Identify Themes," *Field Methods*, Vol. 15, No. 1, 2003.

UNESCO—*See* United Nations Educational, Scientific and Cultural Organization.

United Nations Educational, Scientific and Cultural Organization, Institute for Statistics, "New Methodology Shows That 258 Million Children, Adolescents and Youth Are Out of School," fact sheet, UIS/2019/ED/FS/56, September 2019. As of April 7, 2023:
http://uis.unesco.org/sites/default/files/documents/new-methodology-shows-258-million-children-adolescents-and-youth-are-out-school.pdf

Wetheridge, Louise, "Girls and Women's Literacy with a Lifelong Learning Perspective: Issues, Trends and Implications for Sustainable Development Goals," briefing slides, United Nations Educational, Scientific and Cultural Organization, 2016. As of April 7, 2023:
https://unesdoc.unesco.org/ark:/48223/pf0000244959

Wolfe, Rebecca, Elaine Lin Wang, and Benjamin K. Master, *Developing Enabling Environments for Women's Access to Education and Vocational Opportunities: A Case Study of Policy Change Efforts Led by UN Women's Second Chance Education Programme in India*, BHP Foundation, RR-A259-5, 2023a.

Wolfe, Rebecca, Elaine Lin Wang, and Benjamin K. Master, *Engaging a Global Network to Curate and Disseminate Evidence About What Works to Improve Family-School Engagement: A Case Study of the Center for Universal Education's Family-School Engagement Playbook Effort*, BHP Foundation, RR-A259-3, 2023b.

Wolfe, Rebecca, Elaine Lin Wang, and Benjamin K. Master, *Taking a Contextualized Approach to Synthesizing and Mobilizing Evidence About What Works to Improve Education: A Case Study of Queen Rania Foundation's Contextualized Teaching and Learning Toolkit Effort*, BHP Foundation, RR-A259-4, 2023c.

World Bank, "Ending Learning Poverty: What Will It Take?" webpage, 2019. As of April 7, 2023:
https://openknowledge.worldbank.org/handle/10986/32553

World Bank, *The COVID-19 Pandemic: Shocks to Education and Policy Responses*, 2020.

World Bank, United Nations Educational, Scientific and Cultural Organization, and United Nations Children's Fund, *COVID-19 Learning Losses: Rebuilding Quality Learning for All in the Middle East and North Africa*, 2021.