Soft Skills, Hard Outcomes:

Emplacing Agency in Girls' Education Practice in East Africa

A Dissertation
Presented to the Faculty of the Graduate School
Of Cornell University
In Partial Fulfillment of the Requirements of the Degree of
Doctor of Philosophy

By
Aubryn Sidle
May 2022

SOFT SKILLS, HARD OUTCOMES: EMPLACING AGENCY IN GIRLS' EDUCATION PRACTICE IN EAST AFRICA

AUBRYN SIDLE, PH.D.

CORNELL UNIVERSITY 2022

ABSTRACT

In this project, I argue that agency is an important soft skill for improving the life and labor market outcomes of vulnerable adolescent girls, and so should be a priority in international education policy and for achieving quality education for all (goal 4 in the United Nations SDG agenda). I argue that the way agency is defined, measured, and practiced should attend to the contextually-specific ways in which agency is constrained, and to the power relationships that have historically privileged international actors over local practitioners, and methodological concerns about internal validity over external validity. These dynamics in the girls' education sector have prevented the emergence of accurate measures and truly scalable approaches to programming. I argue that the practice and measurement agency should be 'emplaced' or situated within the particular economic, cultural, and socio-political milieu that structures girls' lives and opportunities, and I provide examples of how to do this.

In paper 1 I use a review of the education, development, and psychology literature, a benchmarking of girls' life skills curricula, and two empirical case studies to synthesize a working definition of agency and to propose a theory for how agentic capacity is developed through skills-focused or critical pedagogical approaches to programming. In paper 2, I offer an example of how to address the problem of contextually relevant measurement, by developing a measure for adolescent girls' agency in East Africa. The paper utilizes a participatory survey development process to create an 'emplaced' measure of girls' agency as a 4-dimensional

construct made up of both internally and externally facing skills and beliefs. I then conduct an investigation of the survey's construct validity and validity of intended use as a pre-post program evaluation measure—using Confirmatory Factor Analysis (CFA) via structural equation modeling (SEM). In paper 3 I argue that instead of seeking to scale evidence-based interventions, the sector should identify scalable *theories* of practice. I use Group Concept Mapping as a participatory tool for group conceptualization and theory-building, to construct a theory of practice for replicating adolescent girls' agency outcomes—one that is empirically driven and collaboratively built by the expertise of local practitioners and community-driven organizations (CDOs) in East Africa.

BIOGRAPHICAL SKETCH

Aubryn Allyn Sidle is a mixed-methods gender and education researcher who completed her Ph.D. in Development Sociology at Cornell University in 2022. She holds a master's degree in Development Management from the London School of Economics and Political Science and a Bachelor of Arts (magna cum laude) from Cornell University in Anthropology, with a minor in Africana Studies. She is a 2018 Fulbright-Hays DDRA Fellow and the recipient of the 2017 Engaged Cornell Fellowship for translational research. From 2010-2016 Dr. Sidle served as the Executive Director of Advancing Girls Education in Africa (AGE Africa), where she continues to serve on the Board of Directors. Dr. Sidle has spent more than 15 years working in East and Southern Africa consulting on projects related to program design, gender, monitoring and evaluation and the governance of NGOs. Her research on girls' education and girls' agency has been previously published by Michigan State University, the United Nations Girls' Education Initiative, and AMPLIFY Girls.

For Steve, Asha Grace, Benny Jude and Sebastian Amal. You are my heart and soul.

For the many generations and lineages of my family who were denied an education either because of their economic circumstances, their gender, or their place of birth. Especially for my grandparents: my grandmother, Molly, who loved me unconditionally; my Sito, Gloria, who pioneered an associate's degree in the 1940s; and for my Jido, George Peter, who spent his entire life working for, and bettering Cornell University, the institution that he loved, even though he would never have the opportunity to earn a degree.

ACKNOWLEDGEMENTS

I have had the extraordinary privilege of being blessed with a group of academic mentors who were champions of my work and willing to support my unique orientation to our field, and to a variety of methods. First and foremost I would like to thank my Committee Chair Professor Parfait Eloundou-Enyegue whose support and belief in me has allowed me to grow and flourish as a scholar. Thank you for helping me to situate myself and my work in terms of the 'big picture,' for pushing me to be more precise in my thinking, and most of all for seeing and believing in the value of my intellectual contributions to policy and science. I am grateful, especially for your commitment to my unique set of methodological interests—I could not have navigated this journey without your support.

I am also grateful to Professor John Sipple, from whom I have learned both statistics and pedagogy, and with whom I share a passion for education and community-oriented policy making. Thank you for orienting me to a systems-view of communities, and institutions, as a starting point for affecting social change in education. I am equally grateful for the guidance of Professor Lori Leonard, for helping me re-discover my love for qualitative methods and feminist theory. Thank you for always "seeing" completely what I was trying to accomplish and helping me to locate my own guideposts in moments when I was unsure of my direction.

I wish to extend a special thanks to Professor Bill Trochim, who was invaluable for helping me clarify my own epistemological orientation to evaluation. I am privileged to have learned the foundations of evaluation research from you and am especially grateful for your time and attention

in teaching me concept mapping, and for your patience as I grappled with methodological questions in my own time.

Without my passionate colleagues at AMPLIFY Girls and the leaders of each of the AMPLIFY Girls' partner organizations, this research would not have been possible. Your curiosity and trust in me, allowed this work to become reality. I am also grateful to Echidna Giving and Wellspring Advisors for their generous financial support, partnership and willingness to participate in several pieces of this study, and to the Einaudi Center and the AWARE (Advancing Women in Agriculture through Research and Education) project at Cornell University for their generous contributions.

In the past 5 years I have thought often about the phrase "it takes a village," never in my life have I more truly understood the meaning of this phrase, than as a Ph.D. student. The most important member of my village is my husband, Steve, without whom I know with a certainty, this degree would not be possible. He has been my cheer leader, support person, sounding board, and even task master, when I needed it most. He was willing to listen to my wild ideas, help me clarify them, and allowed me to vent my frustrations, pushed me to take time to work even when I was hesitant, and pampered me along the way. His support was unwavering, and his willingness to lead our home lives, without complaint, and without request, is the definition of partnership. From you Steve, I have learned so much about what feminist practice looks and feels like, even as I try to make sense of it through research. Thank you for being my person and my partner.

Thank you, Dad, for being a mirror and sounding board, and for our daily walks to campus—I have cherished this time more than you will ever know. Thank you, Mom, for helping me re-focus

when I needed it, and be true to myself along the way. Thanks to the whole community of people, including my parents, my in-laws (Gerald and Colleen), the other "Grands," and the young adults (especially Laura and Colleen) for being part of the village that loved our children while I was finishing this degree. Thanks also to the friend-family, and especially my sisters: Concepcion, who loves and supports me unconditionally, and Brenda who is both thought-partner and friend in many facets of life and scholarship.

Finally, I am grateful for all those who encouraged me to do a Ph.D. at various stages of my life and career—many of you made the suggestion long before I was ready but saw something in me that I did not yet understand. These folks include, Athman Lali Omar, Jane Fajans, Chet Warzynski, Kevin Coray, Dr. G and Joyce Malombe (who gave me the final push at just the right moment).

TABLE OF CONTENTS

INTRODUCTION

The Ranges of Soft Skills—Identifying the Gap

Soft skills are a growing priority area of research within disciplines interested in improving youth outcomes and yet the term 'soft skills' lacks conceptual clarity and consensus. In education, soft skills are referred to as Social Emotional Learning (SEL) competencies, their scope is well-defined and their benefits understood first and foremost in terms of student learning and schooling outcomes.[1] In Global Development, soft skills first appeared in debates about human capital, employability, and the skills needed to support youths' successful integration in globalized labor markets. For vulnerable populations like women, girls, and youth growing up in resource-limited settings, soft skills are also an implicit part of "empowerment," understood as the capacities, knowledge and skills that support autonomous decision-making.[2] [3]

Whereas economists have produced the most evidence on the consequences of soft skills for youth outcomes, they have also spent the least amount of time on definition and accurate measurement. Consequently, their causal evaluations tell us little about which skills produce the demonstrated effects. While theories in education and global development have offered more robust conceptualizations of soft skill domains and related measures, these have been largely developed in disciplinary silos with the former focused on defining and measuring skills with respect to learning outcomes, and the latter on goals such as 'empowerment' or economic security.

[1] Zins, J. E., Weissberg, R. P., Wang, M. C., & Walberg, H. J. (2004). *Building Academic Success on Social and Emotional Learning: What Does the Research Say?* Teachers College Press.

[2] Kabeer, N. (1999). Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment. *Development and Change*, *30*(3), 435–464. https://doi.org/10.1111/1467-7660.00125

[3] Donald, A., Koolwal, G., Annan, J., Falb, K., & Goldstein, M. (2020). Measuring women's agency. *Feminist Economics*, *26*(3), 200–226.

This lack of coordination between disciplines is problematic for practice: it hinders both consensus on the full domain of soft skills and precise understanding of which skills matter for which youth outcomes. Even more importantly for the purpose of this project, it obscures contextual understanding of which soft skills matter for particular populations and settings. My work begins to address these gaps for practice through an investigation of adolescent girls' agency in East Africa.

My Journey to Questions About Agency

Over fifteen years ago, I sat in a classroom full of teen-agers in Richmond California, as they learned to edit video clips of themselves making public service announcements about the dangerous drugs that, at the time, were invading their schools and neighborhoods. Learning to be actors, video editors, and producers was not in my curriculum. What was in the curriculum was a youth-driven directive: to allow students to decide on their own community service projects and to support them with the necessary resources and skills to execute those projects. I watched students learn extraordinary, non-academic, lessons in that program, and for years remained in touch with many of them and watched their lives unfold. It was clear to me that "agency," or the ability to act to affect change in your own life or community, was an important factor in students' life outcomes, and more importantly in how they *felt* about their lives and futures.

On the forefront of the "youth development" movement in the US, my program in Richmond was experimenting with an "asset building" curricular design for improving youth outcomes, that today is well understood as highly effective for serving young people growing up in resource-limited

and challenging contexts.[4] This early lesson about the value of youth agency was repeated to me again and again by practitioners in very diverse settings, throughout my 10 year career working as an executive and a consultant for girl-serving organizations in East and Southern Africa. Although practitioners in different communities didn't always use the word 'agency,' they frequently described the difference between a young woman who had simply finished school and a young woman who had a vision for her life and community and the ability to pursue that vision and thus leverage the benefits of her education to affect change in her life.

This suggested to me that the ability to 'affect change' is similarly important but perhaps conceptualized differently around the world. What does it mean to have agency if you are a teen-ager living in Richmond California, at the height of a gun and drug epidemic, versus a teen-ager growing up in a high-density urban settlement in Kenya, versus a teen-ager growing up in a rural farming community in Malawi? Furthermore, what does it mean to have agency in any of these contexts especially if you are female, or belong to religious, ethnic or racial minority populations? If the work of intersectional feminist theorists (like bell hooks and Kimberlee Crenshaw) has taught us anything, surely the capacities required to affect meaningful change in our own lives vary with our cultural, social, economic, and political milieu.

Furthermore, I noticed with concern that the programs seemingly doing the most for youth outcomes (evidenced by extraordinarily high graduation or transition rates, low rates of teen-pregnancies in communities where high rates are a norm or increasing household income) had few ways of demonstrating and communicating the importance of their soft skill programming. This

[4] Lerner, Richard M. & Benson, Peter L., eds. (2003) *Developmental Assets and Asset-Building Communities: Implications for Research, Policy and Practice.* Springer, Boston MA.

has contributed to a reality where practitioners, donors and policy makers alike, acknowledge that capabilities related to agency are critical for young peoples' success, but little financial investment is made in programs that support these capacities because impact is hard to demonstrate.

How Global Education Policy Came to Soft Skills:

The importance of soft skills seemed to capture the imagination of the global education community in a widely-read report published in 2013 by The Learning Metrics Task Force—a multi-stakeholder Task Force Commissioned by the UN Secretary General's Special Envoy for Education (Gordon Brown) and UNESCO, housed at the Brooking Institution's Center for Universal Education. The report was an attempt answer why, despite substantial progress towards the Millennium Development Goal of universal access to primary education, millions of students, especially in sub-Saharan Africa, were not learning,[5] and why girls specifically were still under-enrolled in school compared to boys in at least 80 countries.[6]

The Learning Metrics Task Force identified seven domains of learning for "quality education" including at least three domains that included soft skills (e.g. social and emotional learning, learning approaches, and aspects of culture and the arts),[7] launching a portfolio of work focused on the "breadth of skills" required by young people for quality educational and life outcomes.[8] These policy reports were meant to influence the framing of the new Sustainable Development

[5] Jones, S., Schipper, Y., Ruto, S., & Rajani, R. (2014). Can your child read and count? Measuring learning outcomes in East Africa. *Journal of African Economies*, *23*(5), 643–672.

[6] King, E. M., & Winthrop, R. (2015). *Today's Challenges for Girls' Education* (Working Paper Working Paper 90; Global Economy and Development at Brookings). The Brookings Institution.

[7] Learning Metrics Task Force. (2013). *Toward Universal Learning: Recommendations from the Learning Metrics Task Force Summary Report* (Learning Metrics Task Force). UNESCO & The Brookings Institution.

[8] Care, E., Anderson, K., & Kim, H. (2016). *Visualizing the Breadth of Skills Movement Across Education Systems* (Skills for a Changing World). The Brookings Institution, Center for Universal Education.

Goals (SDG) agenda for education by expanding the definition of 'quality education' to include a broader view of skills, values, and attributes that support student success. Despite what seemed like global consensus on the need for soft skills in the SDG learning agenda, SDG 4-"Quality Learning for All"-includes no specific indicators related to soft skills, but merely nods to them briefly through targets related to 'skills for promoting global citizenship and sustainable education.'[9]

The Omission of Soft Skills as a Problem of Measurement

This omission of soft skills from SDG4 is glaring, given the previous attention invested in highlighting the importance of non-academic skills. It is also unsurprising, given the lack of widely accepted quantitative measures of soft skills and the inability to define 'soft skills' for a global population. In an attempt to fill this gap, several labs have been dedicated to the definition and/or measurement of soft skills domains since 2015, including the EASEL lab at Harvard University, Skills for a Changing World project at Brookings Institution and the ALiVe Project in East Africa.[10] The EASEL lab's taxonomy project, for example, is consolidating the various frameworks that exist for defining SEL in programs around the world,[11] while Brooking's portfolio is focused on consolidating tools and evidence for measuring 'global citizenship'[12] and '21st Century Skills.'[13] Of these efforts, only ALiVe has offered a substantive contribution to

[9] Care, E., & Anderson, K. (2016). *How Education Systems Approach Breadth of Skills* (Skills for a Changing World). The Brookings Institution, Center for Universal Education.

[10] Mugo, J. (2020). Welcome ALiVE—RELI's Firstborn Initiative (Assessment of Life Skills and Values in East Africa). Regional Education Learning Initiative. https://reliafrica.org/welcome-alive-relis-firstborn-initiative/

[11] Brush, K. E., Jones, S. M., Bailey, R., Nelson, B., Raisch, N., & Meland, E. (2022). Social and Emotional Learning: From Conceptualization to Practical Application in a Global Context. In *Life Skills Education for Youth* (pp. 43–71). Springer.

[12] The Center for Universal Education. (2017) Measuring Global Citizenship Education, A Collection of Practices and Tools. The Brookings Institution, Washington DC. https://www.brookings.edu/wp-content/uploads/2017/04/global_20170411_measuring-global-citizenship.pdf

[13] Care, E., Griffin, P. & Wilson, M. eds. (2018) *Assessment and Teaching of 21st Century Skills.* Springer, Cham. https://doi.org/10.1007/978-3-319-65368-6

conceptualizing the variation of soft skills across place and none have so far produced robust measures of skills.

The Problem of Scale

> "The formal education system offers a strong platform for life skills development. One key advantage is the ability to develop skills continuously alongside academic learning, building on previously cultivated skills as youth graduate to higher grades. Youth can also be targeted at scale and across all ages." (Sridharan et al. 2019, p. 44)

The conclusion of many research institutes, think tanks, and policy makers is that the answer to this challenge of soft skills is to identify scalable programs/curricula that deliver soft skills and scalable measures to evaluate these programs that might be taken up by national education systems.[14] Scalability, however, assumes that soft skills are equally important in every locale, and are likely to be cultivated by the same curriculum in every community and country. Given what we understand about the diversity of socio-cultural, political, and economic contexts, this is unlikely to be true.

Nonetheless, the past decade has yielded a body of research, led by international NGOs with the technical expertise and resources to conduct large-scale experimental evaluations that demonstrate the causal effect of multi-component girls' education programs as evidence for their presumed scalability. Both bilateral and private donors, desperate for investment opportunities to support the achievement of SDG4, have mistakenly equated impact for generalizability (a problem I address

[14] Sridharan, S., Ravindranath, P., Pottinger, E., & Cosentino, C. (2019). *Building Youth Life Skills: Lessons Learned on How to Design, Implement, Assess and Scale Successful Programming*. The Partnership to Strengthen Innovation and Practice in Secondary Education (PSIPSE) & Mathematica. https://drive.google.com/file/d/1-kV-XHhoxMUz1jyTTFgaRPK7eaE9ouO8/view

in paper 3), investing millions since 2015 in INGOs like Room to Read, FHI360, Plan International, Camfed, Mercy Corps, BRAC, and Save the Children,[15] whose large infrastructures promise opportunities for scale in education. Even though these impact evaluations rarely use precise or comparable measures of soft skills, let alone offer robust data on how or why specific program interventions worked to achieve particular outcomes, they have largely been taken whole-sale by the global development community as evidence of scalability.

As a result, small organizations, Southern-led organizations, organizations with single-country or single-region foci and community-based organizations have largely been left out of the global policy conversations on soft skills because they do not possess the financial resources or technical capacity to conduct such evaluations nor infrastructure to scale internationally. Consequently, local innovation and local voices on the question of soft skills and context have been largely ignored.

Motivation for Research

I came to this project as a scholar-practitioner, seeking to understand the soft skill capacity of 'change making' or 'agency'—how it can be developed in young people and the implications of developing this capacity for their life and education. I focus on adolescent girls because their lives are uniquely and doubly structured by the norms that affect adult women as well as the norms that affect all children in their respective communities.[16] In many settings, girls' rights and opportunities are viewed as unique from boys, but so too is their power and potential for economic

[15] These included Michelle Obama's Let Girls Learn Initiative, DFID's Girls Education Challenge Fund (GEC) and the multi-donor collaboration Partnership for Strengthening Innovation and Practice in Secondary Education (PSIPSE).

[16] Mensch, B. S., Bruce, J., & Greene, M. E. (1998). *The Uncharted Passage: Girls' Adolescence in the Developing World*. The Population Council.

and social change.[17] My project centers the voices of community practitioners who are serving under-resourced girls in East Africa.

Structure of the Dissertation

For the empirical side of my project, I collaborate with AMPLIFY Girls (AG)—a collective of 18 community-driven organizations located in rural/urban Kenya, Tanzania, Rwanda, and Uganda. Central to AG's mission is the desire to shift power and resources from global development actors to local communities and organizations. In that pursuit, AMPLIFY Girls and I share a common goal. Accordingly, this project engages with quantitative methods as the language of political power but seeks to repurpose that language in pursuit of 'localizing' the conversation. During the course of my project, AMPLIFY Girls went from having a single employee and founder (based in the US) to having four fulltime employees (three in in Nairobi) governed by a Board of Directors elected from the membership.

I was initially invited by AMPLIFY Girls to lead a process for developing shared metrics that would allow the network to document its collective impact on girls throughout the region. Each AG member organization operates in a distinct community and offers a unique set of services and programming. The process of developing these shared metrics revealed that, despite this diversity of context and programming, AG members also cared deeply about agency as a priority outcome for girls. "Girls' agency and leadership" (as defined by AG members) was one of four shared outcomes produced by the collective and was the only outcome that was unanimously supported by each of the 18 member organizations.

[17] Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development*, *22*(9), 1325–1343.

Below is a description of the three papers of my dissertation and some priorities coming out of these papers for future investigation. Like my journey towards agency, much of my learning has been experiential, and the progression of my thinking on the subject of agency is reflected in the sequence of these three papers. I briefly summarize learning at the end in a short conclusion.

Paper 1 (Define): "Action on Agency—A theoretical Framework for Defining and Operationalizing Agency in Girls' Life Skills Programs," is a largely theoretical attempt to understand and define *agency* from the literature on human development and psychology, education, and gender and development. I construct a hypothesis about *how* programs develop agency in adolescent girls by reviewing some of the practice literature and examining two different practice case studies of critical pedagogy in India and Mexico. It contains my earliest thinking on agency. A version of this paper was published in the Gendered Perspectives on Development Working Paper Series at Michigan State University (2019).

Paper 2 (Measure): "Assessment of a Practitioner-Derived Framework for Measuring Girls' Agency in East Africa," builds on the theoretical work of paper 1 to develop and test a measure of agency for adolescent girls in East Africa. I collaborate with the AMPLIFY network to apply the conceptual framework developed in paper 1 to the East Africa context, and develop a survey instrument to measure girls' Agency regionally. I refine the conceptual framework of agency from paper 1, based on practitioner input, and use this framework to empirically investigate the structure of agency. I conduct a confirmatory factor analysis to confirm the structure of the agency instrument on quantitative data collected from almost 2,000 adolescent girls across 4 countries between Aug 2019 and June 2020. AMPLIFY employee and research collaborator Brenda Oulo

supported the statistical analysis on this paper, which is currently under-going a 2nd review at the *Comparative Education Review*.

Paper 3 (Practice): "Towards a Theory of Best Practice for Girls' Agency in East Africa." Paper 3 is both a theoretical and empirical undertaking. In this paper, I collaborate with AMPLIFY members to build a theory of program practice for producing girls' agency. I do this using Group Concept Mapping as a participatory mixed methodological tool for evaluation and program design.[18] The data used in this analysis was qualitative data collected from AMPLIFY members and representatives from four of AMPLIFY's donors. The paper is intended to be both a methodological contribution to the question of cross-country generalizability as well as a substantive contribution to the question of best practice for improving girls' agency.

There are two areas of follow-up investigation that I would mark as priorities emerging from this research. The first is the need to empirically verify the theory of practice developed in paper 3. As schools and programs reopen in the East Africa region, this might be possible in the years to come through a mixed methods of evaluation of AMPLIFY Girls' programs utilizing longitudinal data on girls' agency and qualitative data on program structure, process, and impact. School and program closures commenced in East Africa along with the rest of the world in March 2020, and although they reopened in Tanzania a few months later, Kenyan and Rwandan schools remained closed for nearly a year, and Ugandan schools were closed for two years—the longest school closures in the world.

[18] Trochim, W., & Kane, M. (2005). Concept Mapping: An introduction to structured conceptualization in health care. *International Journal for Quality in Health Care*, *17*(3), 187–191.

The second priority is a qualitative investigation of how girls themselves view agency and the variation of those views between countries and communities. I conducted informal interviews with girls in four countries over the course of the project and what I learned in those conversations has influenced my writing of papers 2 and 3, but more formal qualitative study with girls themselves would certainly help to clarify the conceptualization, measurement, and development of program practice around agency in the region.

PAPER#1: ACTION ON AGENCY—A THEORETICAL FRAMEWORK FOR DEFINING AND OPERATIONALIZING AGENCY IN GIRLS' LIFE SKILLS PROGRAMS

Originally published in the *Gendered Perspectives in Development* Working Paper Series at Michigan State University's Center for Gender in Global Context in February 2019. (Minor revisions made)

Aubryn Allyn Sidle

Abstract

Debates in girls' education policy and practice point to life skills programs as a promising intervention for improving the outcomes of marginalized girls. Yet the difficulties of defining and measuring life skills and the complexities for practice of understanding how to instill soft capacities like self-efficacy which are often the focus of life skills programs, have contributed to serious knowledge gaps. This article synthesizes theories from education, developmental psychology, and gender and development to offer agency as a locally adaptable framework for the measurement of life skills programs. Based on empirical examples and curricula from five life skills programs, I propose that agentic capacity is critically linked to identity formation processes in adolescence and that alternative pedagogical practice and skills-based learning are important facilitators of agency formation. I offer a draft theoretical framework for how agentic capacity can be cultivated by girls' education programs.

Introduction

The Promise of Girls' Education

Evidence from around the world shows that education continues to have a powerful effect on reducing global inequality (Giroux et al. 2020). Specifically, increasing girls' schooling dramatically reduces child marriage (Wodon et al. 2018), teen fertility (Eloundou-Enyegue & Stokes 2004), maternal death (Bhalotra et al. 2013; Gupta et al. 2002), and child mortality (UNICEF 2010), while increasing national GDP (Klasen 2002). Recently, however, questions about gross inequities in school quality (Jones et al. 2014, Learning Metrics Task Force 2013) have challenged the traditional assumptions about the implicit value of education, especially for girls (King & Winthrop 2015).

Differential experiences of education for boys and girls (Mensch et al. 2003) combined with real concerns about school safety for girls (Psaki et al. 2017), have led to a growing acknowledgment that education that does not actively address harmful gender norms, can be disempowering (Murphy-Graham & Lloyd 2015; Adely 2012; Dehyle 2009). Yet the promise of an "empowering" and "quality" education, continues to have powerful implications for gender equity (Sahni 2016), social mobility, and the positive transformation of society (Sperling &Winthrop 2016). The link between girls' body autonomy, sexual and reproductive health, and education, make the gains from girls' schooling most profound during adolescence, thus amplifying the benefits of girls' education at the post-primary level (Wodon et al. 2017; Sperling & Winthrop 2016).

Girls' Life Skills Programs for Empowering Education

In this paper, I take the notion of an empowering education as the goal of education efforts globally and seek to understand and untangle the circumstances of practice which might contribute to education becoming an enabling force for girls and fulfilling its promise of opportunity. The overarching question of this study is: How can practitioners ensure that girls stay in school through secondary and can translate their educations into viable change? Recent conversations in policy and girls' education practice point to informal life skills education programs as successful for both improving girls' retention and achievement in school, as well as enhancing the effects of girls' schooling, such as improving health and labor market outcomes (Dupuy et al. 2018; Kwauk & Braga 2017).

Specifically, life skills programs have shown marked success at supporting girls' own abilities to address unhealthy relationships between gender and power (Haberland 2015) and support the development of soft skill capacities such as self-confidence, self-awareness, leadership, and self-

efficacy (Temin & Heck 2020; Dupuy et. al. 2018, Austrian et al. 2016). Yet, the diverse and often disparate range of skills included in life skills programs, as well as the qualitative nature of many of the skills enshrined in life skills curricula, begs an obvious set of questions: What are life skills? How can they be measured? And which skills are important for which girls? To date there is no sector-wide consensus on these issues to inform practice, policy and investment. In fact, successful life skills programs cover a wide array of skills and competencies including (but not limited to), sexual and reproductive health knowledge, social and interpersonal skills, self-awareness, job readiness training, leadership, entrepreneurship, financial literacy, and gender rights (Dupuy et al. 2018).

Agency as a Causal Mechanism of Life Skills Programs

The rise in academic interest in life skills and by extension, agency, has been driven by the global focus on redefining the parameters of a "quality education" to include a variety of soft or "21st Century Skills" that show promise for later life outcomes (Winthrop and McGiveny 2016). Driven by revelations about reduced student learning despite increased enrollment globally, soft skills were clearly identified as part of an expanded definition of "quality education," by a policy task force charged with informing the sustainable development agenda for education (Learning Metrics Task Force 2013).

Subsequent evaluation research suggests that agency is particularly important soft skill for marginalized girls' education outcome in countries where access to goods, services, and resources of any kind come at a high premium and where the most valued role for adolescent women is as a wife and soon-to-be mother (Murphy-Graham and Leal 2015; Sahni 2017; Santhya et al. 2015; Sperling and Winthrop 2016; Warner et al. 2014). A recent report published by the Brookings Institution identifies agency as one of two critically important mediating

factors required for girls to translate the knowledge, skills, and attitudes (KSAs) gained in educational programs into empowered action (Kwauk and Braga, 2017). In this paper, I propose that agency is not separate from the knowledge, skills, and attitudes cultivated by girls' life skills programs, but rather is the desired outcome of the right combination of these competencies. In other words, the success of life skills programs at improving the education, health, and economic trajectories of girls is a result of their ability to improve girls' agency.

Yet agency, much like girls' own lives and cultures, is not a monolithic construct. I argue that the vast majority of life skills programs attempt to instill a set of positive self-beliefs and personal competencies in young women, and that together, these competencies and self-beliefs create agency or agentic capacity.[19] I hypothesize that the means towards agency (the specific skills and self-beliefs required to foster agency) is likely to vary across place, based on the specificities of opportunity structures, culture, and context. Whereas some skills may be universally important for girls' success across places, practitioners are now beginning to acknowledge that there are likely a host of skills and capacities that can and should be locally defined (Dupuy et al. 2018). Thus the current diversity of life skills curricula can be seen as a reflection of the diversity of contexts in which girls live. I address the issue of how to define and evaluate life skills programs by offering a culturally and locally adaptable definition of agency as a framework for measuring life skills programs across place.

[19] Agency and agentic capacity are used somewhat interchangeably throughout this paper. However, there is a subtle distinction. Agency is the set of skills/self-beliefs required to exercise "agentic capacity." Agentic capacity being the actual ability to exercise agency. The author assumes that if one has Agency, then one also possesses agentic capacity.

The rest of this paper is organized as follows: First, I review definitions of agency from across the social sciences literatures and synthesize an operational definition of localized agency. Second, I offer a brief review of life skills curricula and the evidence on agency as it relates to girls, women, and education showing that agency as an organizing frame for life skills is bolstered by a surprising consensus across disciplines on the importance of Agency to education. I then discuss two empirical examples of girls' education programs that investigate promising pedagogical strategies developing agency for girls and I use lessons from these examples to offer a draft theory for how life skills programs can successfully cultivate agency.

Towards a Definition of [A]gency

Coming to a clear understanding of agency as a key outcome is complicated by the fact that, much like life skills, it lacks a common definition. In the social science literature, agency is most commonly utilized to denote any kind of choice or action taken by an individual (Kabeer 1999, 438), and operationalized for measurement purposes as decision-making (Donald et al. 2017). However, the type of agency referred to in this paper is much more specific than this broad notion of exercising choice. The agency herein refers to a particular capacity—an ability to coordinate multiple skills, resources, and contexts to achieve a desired strategic outcome. It is an understanding of agency that looks at the internal capacity of an individual and her relationship to power—both her own internal power to take control over her life, but also her ability to navigate the external power structures that may either constrain or enable her actions. This type of agency is one we might distinguish with a capital 'A' and is concerned with the skills required to achieve a desired goal within the specific socio-cultural environment and network of power relations in which an individual is embedded.

In the introduction to her book *The Structure and Agency of Women's Education*, Mary Ann Maslak offers two definitions of agency that are roughly parallel to the distinction made in this paper. First, Maslak identifies *oppositional agency*, or agency that is actively challenging the "established norms in a system" to achieve a desired outcome, and juxtaposes this to what she calls *allegiant agency,* which refers to the actions of an individual in making choices that "align with popular thought" (2008, xv). Maslak's definition is dynamic because it places at the center the *structures* of constraint—what she calls the "established norms"—within which individuals must choose and act. It is similar to Naila Kabeer's definition of agency as the "power-to" or, more specifically as "people's capacity to make their own life choices and define their own goals, even in the face of opposition from others" (1999, 438). Also explicit in Kabeer's construction of agency is the idea of goal-setting and action towards those goals. In a review of metrics used to measure women's agency around the world, Donald et al. expands Kabeer's definition to include three core capacities:

1. Ability to define goals "that are in line with their values;"
2. Self-perceived control/ability; and
3. Action towards goal achievement (2017, 6).

The specification of goal-setting as a skill that must be guided by an individuals' own values is a critical distinction and one that mirrors Maslak's oppositional and allegiant categories. Implicit in the idea of aspirational goal setting is the need for positive self-belief. In other words, a person must believe that she is capable of achieving her goals before achievement is possible.

The current consensus in development on goal setting and goal achievement as central to agency, glosses over the importance of specific skills beyond goal setting that might be important to strategic action—things like strategic thinking, inter-personal communication, decision making,

self-awareness, etc. Current conceptualizations of agency also do not distinguish between action and effective action, nor do they flush out the kinds of positive self-perceptions that are likely required for effective action towards aspirational goal achievement. A more detailed, and nuanced approach to defining agency is necessary for understanding how agency might be measured, how life skills programs should be designed to cultivate agency and the variation of skills needed to produce agency based on various structural constraints (Maslak 2005) or oppositional forces (Kabeer 1999).

For this type of specificity, we must turn to the field of developmental psychology, which offers some compelling insight into how agency operates. Among the earliest writers on the subject of agency within this field was Fritz Heider who conceived of agency as an "effective personal force," which he believed was "underpinned by a combination of power or personal ability, the intentions to try, and any relevant environmental factors" (Little et al. 2006, 64). Herein we see Heider offer a rough framework for understanding agentic capacity as one that connects skills or abilities to intention and the structure of the environment. Heider's definition offers a bit more direction for practitioners because it differentiates between aspects of agency which can be explicitly taught by written curricula in more or less traditional settings (skills or abilities), versus aspects which are dependent on personal development or environmental constraints (intentions and structure).

Albert Bandura adds to Heider's view by identifying the self-efficacy capacity of human agency as a necessary pre-cursor for developing agentic ability (1982, 1989, 1993). Bandura divides agency into two functions—perceptions or self-beliefs (in lieu of Heider's intention to try) and

coordinating skills or abilities. For him, the most important aspect of human agency is self-efficacy or the self-belief that a person "can produce the desired effects and forestall undesirable ones by their actions" (Bandura 2009, 179). Bandura defines agency skills as the actual ability to organize self-beliefs like self-efficacy, actions and resources to "influence intentionally one's functioning and life circumstance" (2006, 1).

On the subject of environment, Bandura provocatively notes that although individuals act as their own agents in the world, they also act as agents of their environment, thus implying action is constrained by structural forces (1989, 1181)—a point also acknowledged by Maslak and Kabeer. However, none of these scholars discuss how agency might also be *enabled* by environmental factors and how environment may, itself, be influenced by individuals' agency—a notion popularly conceptualized by psychologist Urie Bronfenbrenner whose bioecological theory of human development posited that youth develop as a process through which they shape and are shaped by their environments (Bronfenbrenner & Morris 1994). When it comes to girls' education practice, questions about environmental influence are critical as life skills programs grapple with the balance of interventions that focus on individual skills and capacities versus the opportunity structures and environmental constraints that necessarily influence outcomes (Kwauk and Braga 2017).

Little et al. (2006) expand Bandura's theory of human agency placing it within a broader theory of the development of human agentic capacities throughout the life span, noting that agency is comprised of two different types of capacities—causal capability and agentic capability—the former being defined as the "knowledge, behavioral skills, self-perceptions and beliefs about

one's environment" required to demonstrate the latter (67-68). Within Little et al.'s framework a number of skills commonly associated with life skills in the education and development literatures are identified as examples of either causal or agentic capacities. These include skills such as self-awareness, self-regulation, goal-setting, problem solving, and decision-making (ibid).

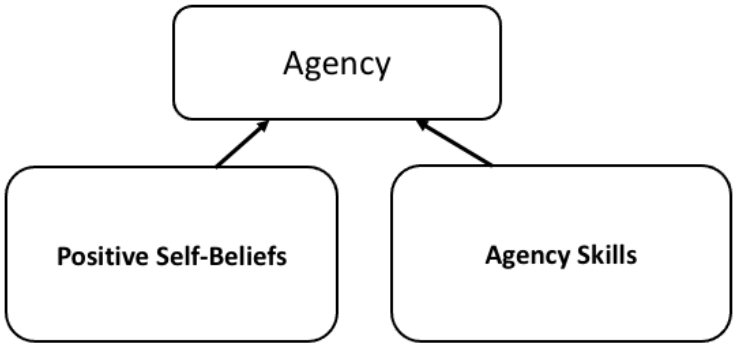
Although, these theories of agency are strongly rooted in Western notions about autonomy, individualism, and personal desire, they still offer a foundation for understanding alternate or culturally-mediated forms of agency (Ahearn 2010), and what those might look like when culturally emplaced. In sub-Saharan Africa for example, rapid urbanization, demographic shifts, and the related changes to family structure, have introduced new social, political, and economic contexts in which opportunities are accessed based on new and more fluid notions of culture, tradition, and community (Giroux et al. 2010). Defining a flexible framework for understanding agency as it is currently being deployed in practice, actually allows for local negotiation, interpretation, and understanding in ways that can meaningfully improve practice.

From Heider, Bandura, and Little et al., the distinction made between the skills and self-perceptions function of Agency is particularly useful for practice and reflective of the current field. Bringing together psychological theories and evidence from development and education practice, I offer my own synthesis and working definition of "big A" Agency—one that can be conceptually interrogated and operationally deployed (see Figure 1):

> **[A]gency**—The capacity of individuals to define aspirational goals[20] and coordinate the knowledge, skills, attitudes[21] and resources both internally available to them (individual capacities) and externally available to them (in their social, institutional or physical environments) in order to take action to achieve stated goals.[22]

Thus, Agentic capacity is made up of both **positive self-beliefs** (including self-efficacy but not excluding things like self-esteem and self-confidence), and concrete **skills** which can be learned over time.[23] These skills include core information or knowledge as well as internally and externally-facing skills that are required for goal-setting and achievement.

Figure 1: A working Definition of [A]gency



At a minimum, this definition of Agency provides practitioners with a framework for housing and thinking about many of the individual skills and attitudes that are already the focus of life skills programs in girls' education and that are at the center of the global citizenship skills conversation under the new SDG framework. Often specific soft skills (e.g., goal setting) and attitudes (e.g., self-confidence) are discussed in silos—disconnected from a theoretical

[20] (Donald et al. 2017)
[21] (Kwauk & Braga 2017)
[22] Based on Albert Bandura's notion of agency but expanded to include a broader definition of resources
[23] Bandura (2009) theorized the importance of self-efficacy in agency, while Little et al. (2006) developed a theory of agentic capacity which included skills and self-perceptions.

framework that helps practitioners to understand that skills and attitudes are interconnected and important "package deals" in programs that seek to support girls' Agentic development.

Given what we know from both the theoretical review above, and the forthcoming case studies (below), I agree with Bandura that self-efficacy—the belief in one's own ability to achieve stated goals—is among the most influential self-perceptions involved in the human Agency function and is a necessary pre-requisite for Agentic capacity (Bandura 1989; Donald et al 2017). Similarly, given their prevalence in the literature we might hypothesize that specific skills such as goal setting, decision-making, self-awareness, and inter/intrapersonal communication are among the most universally important skills that make up Agentic capacity, although these are likely to change based on the environmental and social constraints of each place. There is a clear need for practitioners to test and revise the specific skills, self-beliefs, and their combinations that might be most relevant for developing Agency in particular communities, to better understand which combinations are most important for which contexts.

Agency in the Practice Literature

Agency in Life Skills Education Programs

When we consider the specific curricula of life skills programs from across the field of practice, we can see this delineation between positive self-perceptions and skills clearly reflected in both the content of curricula and in the outcomes identified for measurement. Table 1 presents a curricular review of five life skills programs operating across sub-Saharan Africa. These curricula are those that have been made publicly available, have garnered a fair amount of acclaim for their quality through quality evaluations and/or randomized control trials, as well as citation in systematic reviews. The table organizes curriculum content into skills (both internally- and externally-facing) and positive self-beliefs. The selected curricula are drawn from two

primary types: those that were funded by international development donor communities for use in a diversity of countries and those that were developed by the local leadership of single-country-focused NGOs, and tailored to the communities and populations served.

Table 1: Summary of Five Sub-Saharan Africa Focused Life Skills Curricula

Curriculum	Author	Country	Self-Beliefs	Skills	
				Internal	External
CHATS: Creating Healthy Approaches to Success	AGE Africa 2016	Malawi	Self-efficacy** Self-esteem**	Decision making***	Interpersonal communication**
				Self-awareness***	Public speaking**
				Goal setting***	Negotiating romantic relationships**
				Positive gender attitudes*	Preventing pregnancy**
				Resisting peer-pressure**	Conflict resolution*
				Self-advocacy*	Identifying Gender based violence*
				Leadership*	Negotiating sex and sexuality
				Stress management*	Legal literacy
				Personal hygiene/menstruation	Identifying gender power imbalances
					Group facilitation
Educate Leadership & Entrepreneurship Course	Educate! 2014	Uganda	Self-efficacy**	Decision making***	Public speaking**
				Self-awareness***	Conflict resolution**
				Goal setting***	Interpersonal communication**
				Problem solving	Financial literacy**
					Financial Planning**
					Collaboration/teamwork*
					Resource mobilization
					Peer mentoring
					Networking
Health and Life skills Curriculum	Population Council 2015	Multi-country (piloted in Zambia)	Self-esteem**	Decision making***	Interpersonal communication**
				self-awareness***	Public speaking**
				Goal setting***	Conflict resolution**
				Resisting peer-pressure**	Community service*
				Positive gender attitudes*	Financial planning**
				Flexibility (resilience)*	Financial literacy**
				Leadership*	Negotiating romantic relationships**
				Life planning*	pregnancy prevention**
				Stress management*	Identifying Gender Based Violence*
				Self-advocacy*	Negotiating sex & marriage
Ambassador Girls' Scholarship Program Guide to Mentoring & Life Skills	USAID & AED 2008	Multi-country (Malawi, DRC, Tanzania, Kenya, Ghana)	Self-esteem**	Decision making***	Negotiating romantic relationships**
				Self-awareness***	Preventing pregnancy**
				Goal setting***	
				Resisting peer pressure**	
				Leadership*	
				Decision making***	Interpersonal communication**

Akazi Kanoze, Youth Livelihoods Project	USAID (PEPFAR) & EDC 2010	Multi-country (piloted in Rwanda)	Self-efficacy**	Self-awareness***	Financial literacy**
				Goal setting***	Financial planning**
				Life planning*	Collaboration/teamwork*
				Flexibility (resilience)*	Job Seeking skills
				Exercising rights	
*** Present in 5 out of 5 reviewed curricula ** Present in 3 or more curricula * present in 2 curricula					

From the short list in Table 1, we see a few key patterns emerge. First, although there is a broad range of skills covered in these curricula (from financial literacy to negotiating healthy romantic relationships), there is also a fair amount of overlap—particularly on which positive self-beliefs are covered and internally-facing skills. Three out of five curricula have content area explicitly focused on teaching self-efficacy and self-esteem. Notably, both locally focused curricula identified self-efficacy as central—in agreement with Bandura, Little, and the other theorists referenced in the previous section.

Second, there is near consensus on a set of internally-facing skills with 4 out of 5 curricula adopting decision-making, self-awareness and goal setting—which is also a clear reflection of what we've learned from our literature review to-date. Third, there is much less of a consensus across curricula on the externally-facing skills that are important—with no skills being represented in even 4 out of 5 of the curricula reviewed and only a few appearing in at least 3.

Such a diversity of externally-facing skills supports the argument that as social, cultural, and economic contexts change, so, too, does the set of skills needed for improved outcomes. Agency, as it is presented in this paper, allows for this diversity, while still providing a comparable yardstick for measurement. In other words, under our definition of Agency, life skills programs that focus on financial literacy, and financial management skills and those that focus on

negotiating romantic relationships, for example, can both be considered as contributing to improved Agentic capacity.

In Dupuy et al.'s review of 12 informal life skills programs' key outcomes, 10 out of 14 identified the concrete skills (e.g., decision-making, goal setting, financial literacy, etc.) and knowledge content (e.g., sexual and reproductive health) as key outcomes, and eight out of fourteen evaluated positive self-perceptions (e.g., self-esteem, self-efficacy, or self-confidence) (2018). Those that didn't identify positive self-perceptions as a key outcome identified these as one of several goals for the program that were presumably unmeasured due to the difficulty of evaluating such non-cognitive capacities.

Agency as a Skill of Interest Across Disciplines

Bolstering the argument for Agency as an organizing frame for life skills programs is both empirical and theoretical evidence from a variety of disciplines and methodological practices across the social sciences. These practices have shown that core social-emotional learning competencies, like self-confidence, have contributed to improved academic achievement (Borghans et al. 2008; Gutman and Schoon 2013; Hawkins 1997; Malecki and Elliot 2002; Welsh et al. 2001; Zins et al. 2004). When we narrow the focus of inquiry specifically to core skills related to Agency as defined above, and positive self-beliefs, a surprising consensus exists on the importance of these beliefs and competencies in both education and later life outcomes.

In the gender and development sector Naila Kabeer has established women's agency as a foundational skill for empowerment. Her work has heavily contributed to the notion that such successful interventions for women must consider a broad set of internal capacities such as agency as part of what enables the material circumstances of success. Murphy-Graham and

Lloyd, in their framework for understanding the core competencies and conditions that contribute to an empowering education, similarly identify agency and personal efficacy as two of the necessary pre-conditions that generate the circumstances of empowerment through educational settings (2016, 558). They further divide the core competencies into knowledge, personal competencies, social competencies, and productive competencies. This is a more nuanced and detailed way of thinking about internally-facing skills (e.g., personal competencies and knowledge) and externally-facing skills (social competencies and productive competencies) as I have roughly laid them out in Table 1.

Reinforcing the centrality of agency to women's empowerment programming, qualitative evidence from the fields of anthropology and sociology identify a variety of positive self-perceptions as influential components of the educational experiences of girls from around the world (Adely 2012; Cervantes-Soon 2016; Deyhle 2009; LaBennet 2011; Maslak 2008; Mensch et al. 2003; Miron and Lauria 1998; Sahni 2017; van Santen 2010). In these studies, the key to girls' educational success is a sense of autonomy, self-efficacy, self-confidence, and other aspects of positive self-belief that are produced by enabling environments and pedagogies (Cervantes-Soon 2016; Maslak 2008; Sahni 2017; van Santen 2010). Empirical studies like those referenced seem to point to girls' identities—how they are developed, and either reinforced or discouraged by girls' socio-cultural environments—as a critical factor in the kinds of change produced (ibid).

Causal evidence, primarily from the field of economics, corroborates these findings. Bowles et al. (2001) and Osborne-Groves (2005) conducted studies that showed the positive returns on the

labor market to self-efficacy, while Pamela Lenton's work, causally linked self-efficacy in women to higher educational attainment and earnings (2014). Ibarrarán et al.'s study of a life skills program in the Dominican Republic established a causal relationship between higher levels of hope for the future (aspirational goals), persistence, conflict resolution and leadership skills, with reduced pregnancy for girls and increased labor market outcomes (2014). Although the specific soft skills of interest were not teased out, Furnham et al. (2009) and Lounsberry et al. (2004) show that non-cognitive[24] (soft) skills have a positive effect on reducing absenteeism and dropout rates and increasing educational attainment. These studies build on the work of economists Raj Chetty (2011) and Kirabo Jackson (2012), who established that teachers' ability to improve student non-cognitive skills accounted for much of the variability in longer-run outcomes, and, in some cases, more so than cognitive skills alone.

Promising Approaches for Cultivating Agency

If Agency is, indeed, an appropriate frame for measuring life skills, what do we understand about effectives strategies for developing agentic capacity in girls? From our working definition of Agency, two strategies emerge from the literature of practice as important for the acquisition of agency skills, positive self-perceptions, and the development of Agentic capacity.

The first strategy is the positive youth development approach, which employs skills-based learning and community engagement as tools for improving the developmental assets of youth (Benson et al. 2003). These assets include many of the skills and positive self-perceptions important to Agency, including self-esteem, planning and decision making, and personal power

[24] The term "non-cognitive" was coined by James Heckman (2006) to denote the variety of soft skills and their effects on labor market outcomes.

(Search Institute 1997). Second, programs that utilize critical pedagogy as a strategy for developing youth into agents of change in their own environments offer a powerful model for developing both positive self-perceptions as well as Agency-related skills. The former focuses on skills and helps youth build Agentic capacity and positive self-perceptions through the acquisition of concrete organizing skills and the practice of those skills, whereas the latter pursues a more complex, albeit potentially longer-lasting, strategy of re-educating youth on the structures of power that order their lives and offering them analytical and practical tools for upending those power hierarchies.

Skills-based learning and positive youth development programs have garnered at least two decades of compelling quantitative and qualitative evidence to support their use. Much of this literature is causal and offers a fairly nuanced understanding of particular skills and capacities (Catalano et al. 2019; Alvarado et al. 2107; Larson and Angus 2011; Catalano et al. 2004; Lerner and Lerner 2005a). Among the skills highlighted by Catalano's critical review of more than seventy-seven positive youth development programs, at least one-third have been identified by developmental psychologists as important for agentic capacity: self-efficacy (Bandura 1989), self-determination (ibid.; Little et al. 2006), positive identity formation (Brandtstadter 1999; Larson and Angus 2011; Lerner et al. 2005b), and belief in the future, or hope.

Whereas positive youth development practice has been implemented and tested in many countries, less is known concretely about critical pedagogy for improving youth outcomes. Although strong theoretical arguments support its use, to-date, primarily ethnographic and qualitative empirical evidence exists to promote critical pedagogical approaches in education, all

of which suggest positive effects of these programs on agentic capacity (Cervantes-Soon 2017; Bajaj 2015; Hull et al. 2015; Sahni 2017; Miron and Lauria 1998; Stevens and Slavin 1995). While some positive youth development programs adopt innovative pedagogical approaches in the pursuit of skill development (such as youth-driven, or youth-led program designs), typically these programs do not embrace the complete pedagogical redesign advocated by Paolo Freire (1970), the founder of critical pedagogy, or the lineage of thinkers who followed him (e.g. Henry Giroux, bell hooks, Robert Chambers). Critical pedagogy assumes that endemic to educational practice itself are a set of perverse power relationships—such as that between teacher and student (Giroux 2001) and the various intersections of their identities which must be transformed through a pedagogy that reinvents these power dynamics and places student power at the center of learning (hooks 1994). Critical pedagogy is then, by definition, an effective educational approach for groups of people who experience disproportionate disenfranchisement and oppression, such as the working-class poor; racial, ethnic, and religious minorities; as well as women and girls.

Within the practice of critical pedagogy, learners uplift themselves by becoming aware of the power structures in which they are embedded and by acquiring the skills required to challenge those power structures. This is done through ongoing critical discourse and what Freire called "praxis" or the cycle of discourse, reflection, and action (Freire 1970). In critical educational environments, the primary tool of developing Agency (or "conscientization") is critical dialogue—an active conversation between students and teachers that illuminates the systems of inequality and oppression affecting the lives of students and supports their autonomous development (Freire 1970).

From a bird's-eye view, one might see critical pedagogy as a foundational and pre-cursory practice to a positive youth development approach—one that could be used to uplift and amplify the effects of skills-based learning because skills would be built upon critical social analysis which fosters an empowered self-efficacy. The implication is that skills-based learning, combined with alternative pedagogical approaches which produce self-efficacy, are a powerful combination for cultivating agentic capacity. Indeed, life skills programs have already begun to experiment with alternative or student-centered teaching strategies (Leventhal et al. 2016; Shah 2011). But how does critical pedagogy work to instill positive self-beliefs like self-efficacy? What can we understand about this process? What follows are two examples which attempt to illustrate how critical pedagogical approaches work to foster self-efficacy in education. Both examples are from a school-setting, but their lessons remain relevant for application in formal and non-formal life skills practice. I will take lessons from each to formulate a framework for understanding how agency develops in education practice.

Two Case Studies in Critical Pedagogy from Mexico & India

Claudia Cervantes-Soon's ethnography *Juarez Girls Rising* (2017), offers a detailed analysis of critical pedagogy amongst poor young women attending a public high school in Juarez, Mexico. Altavista is a state-funded secondary school that actively embraces a critical pedagogy, with three tenets at its core:

1. Freedom and autonomy of students;
2. Radical love and reciprocal care; and
3. Critical discourse and activism (133).

Young women's experiences of these principles open a unique and powerful space for self-expression and through it, identity formation or "autogestion" (self-authorship). Cervantes-Soon argues that autogestion is enabled by critical pedagogy which views learning as a collaborative process, where the hierarchy between teacher and student is intentionally broken down. Students are asked to critically evaluate what was being taught to them, examine theories, and offer rebuttals. Despite the many other spheres of their lives where young women's thoughts, appearances, and choices were heavily monitored and policed by the gendered norms of their communities and families, Altavista school is a place where girls' opinions matter, their voices are acknowledged, and where they are actively encouraged to express themselves (144). "Thus young women viewed the freedom that Altavista offered as a precious and unusual opportunity to take control of their own actions and their own learning, to explore the possibilities about who they could be and what they were capable of in a safe space" (143). Thus we see that the environment created by the school's core principle of freedom and autonomy, created the circumstances for positive self-expression:

> The hallmarks of Altavista's critical pedagogy…worked in unison to promote a space of liberation for women and the development of autogestion [self-authorship] as articulated by students and teachers. For the young women…to be a [self-author][25] meant to have the will and embrace the possibility to develop their own mission, to undertake their own project of life...[Self-authorship] was necessary to be able to…pursue their goals with little external support…This is an ultimate form of agency, the most empowering result of the students' education. (177-178)

Agentic capacity at Altavista is cultivated through the careful and painstaking work of critical self-awareness and self-authorship—in other words, the work of identity formation. For Altavista teachers and students, the two are inextricably linked—agency evolves simultaneously with, and

[25] The author utilizes the Spanish word "autogestivas", but I have used the English translation for clarity.

as a result of, students' developing a sense of self and self-determination. Thus critical pedagogy is enabling a space for positive identity formation which in turn builds self-efficacy.

We find similar evidence from India in Uvrashi Sahni's book, *Reaching for the Sky,* which is an ethnographic and autobiographical account of the creation of Prerna School—a primary and secondary girls' school serving the Dalit community in Lucknow, India. Sahni describes the school's pedagogy as a critical feminist one that features critical dialogues as a mainstay of the school's philosophy and success (2017). Sahni defines the academic success of Prerna students as follows:

> To reach the overall goal [of the school], a girl must: Read, write and successfully complete the government mandated syllabus; Learn to recognize herself as an equal person; Emerge with a sense of agency, of control over her life, of aspirations for her future, and have the confidence the skills to realize them; [and finally] Gain a critical understanding of the social and political structures that frame her life and determine its limits and possibilities, which would, in turn, enable her to push the boundaries and reconstruct her life in more empowering ways. (60-61)

Sahni's construction of the methodology of the school is strikingly similar to Altavista—incorporating elements of radical care, critical discourse, and self-authorship. Goal number one of Prerna is identified as building a "universe of care" and a "web of support" that together helps girls to experience what Altavista referred to as radical love from teachers, administrators, and support staff. Much like Altavista, at Prerna this effectively opens up girls' learning by allowing them a safe space for self-expression and creativity. "The key is being attended to and responded to as whole persons not as objects. By confirming each other and rendering each other valuable our lives are made visible and valuable…" (82).

Beginning in grade four, students engage in critical dialogues, where they are given space to share and reflect on their lived realities and are then offered tools to deconstruct these realities and reconstruct them in ways that are more empowering and enabling for their lives (113). Ultimately, however, the goal of Prerna school is to explicitly do by design what Altavista seems to have accomplished implicitly, which is to support girls in the formation of their own unique identities. Sahni places this process as the central goal of education: "The goal of all teaching and learning…is to answer the fundamental question: 'who am I and what is my relationship with the universe and others in it?'" (129). Sahni goes onto articulate how Prerna supports students' development of their own identities by interweaving creativity, performing arts, and critical dialogue throughout the curriculum of the school. Coupled with the structures of care already put in place as a foundation to this learning, Prerna school has extraordinary outcomes—with graduation rates well above the national average and 97% of its graduates continuing on to higher education (194).

Importance of Identity Formation and Enabling Environments to Agentic Development

Although both Prerna Girls and Altavista schools represent formal schooling environments, important lessons about the mechanisms of agentic formation can be gleaned for use in life skills programs which have often experimented with alternative pedagogies. Of the five curricula reviewed in Table 1, at least three utilize alternative pedagogical delivery methods—either peer-led facilitation (AGE Africa CHATS curriculum), peer mentoring (Ambassador Girls' Scholarship Program), or participatory teaching methods (Educate's Socially Responsible Leadership Curriculum and Population Council's Health and Life Skills Curriculum). Several recent evaluations of girls' life skills programs utilize student-centered and student-driven

learning strategies (Leventhal et al. 2016; Shah 2011) and many others where peer-to-peer engagement is a key strategy (Catalano et al. 2019; Austrian et al. 2016; Unterhalter et al. 2014).

From the Prerna and Altavista examples, we can clearly see the importance of identity formation to the creation of positive self-efficacy and the relationship between identity formation and environment. Young women in both these schools are offered an enabling environment in which to explore self-expression and self-concept in ways that had a profound impact on their Agentic capacities. Both Cervantes-Soon and Sahni show how self-expression helped girls to develop self-efficacy and develop personal life goals in accordance with their values. After forming the basis of this self-efficacy—through critical dialogue, community engagement, social justice practice, and activism—girls learned the skills necessary to build self-efficacy into fully-formed Agency.

Returning to our working definition and theory of Agency which conceives of agentic capacity as a combination of skills and self-perceptions, I identify the unique work of critical pedagogy as primarily one that helps to build self-efficacy—an important self-perception necessary for agentic capacity. Secondarily, critical pedagogy seems to support the development of key agency skills such as goal setting and self-awareness. This relationship between identity formation and self-efficacy is also reflected in the development psychology literature. Jochen Brandtstadter believed that self-concept and identity guided and directed the self-regulatory process that is critical to agentic functioning (1999). He identified the relationship between identity development and self-regulation as a critical need of further research and pointed to adolescence as the most important time period for studying self-regulation because of identity formation, and

conceptions of self are among the major developmental milestones occurring at this age, as summarized by Lerner et al. (2005b, 7).

However, such arguments contradict some of the wisdom from global education policy which point to early childhood as the most important time period to begin developing agency and adolescence as potentially too late (Kwauk & Braga 2017, 17). While this might certainly be true for many cognitive skills and competencies, the unique nature of adolescence as a time of the emergent self, when identity and self-concept are the primary task of the brain likely indicate that adolescence is the ideal time to cultivate Agentic capacities, and thus life skills programs are well placed to do so.

Another issue emerges in the untangling of self-efficacy and agentic development—this is one of environment. The Prerna and Altavista case studies succeeded in developing agentic capabilities amongst their students in part because it was first ensured that school environments created an enabling context for the development of agency and transformational learning even when home or social environments did not. This underscores the importance of programmatic environment and perceived safety in that environment—especially when those environments are much less controlled than schools—as is the case for many life skills programs.

Whereas the field of psychology views the development of human capacities as a protracted and teleological process that takes place over the course of a lifetime, identity is viewed from the lens of anthropology as a fluid construct that is constantly developing in relationship to environment (Holland et al. 1998).

> Identity is a concept that figuratively combines the intimate or personal world with the collective space of cultural forms and social relations…we begin with the premise that identities are lived in and through activity and so must be conceptualized as they develop in social practice. (5)

Identity formation—as a process that necessarily captures the social and lived realities of an individual—makes much more sense from the lens of education, where social practice is intimately linked to learning. From this view people can both produce environments and be produced by them (echoing Bronfenbrenner's bioecological theory of human development), indicating human Agency is a powerful factor in mediating lived contexts.

There is both empirical and theoretical evidence to support this, suggesting that Agency itself includes the ability of the individual to influence his or her environment and potentially change the constraining factors that exist (Bandura 1989; Freire 1970; hooks 1994; Larson and Angus 2011; Little et al. 2006; Sahni 2017). A theory of agentic development in life skills practice therefore must look at the specific interaction between individuals and their environments and produce context-inclusive theories about agentic development.

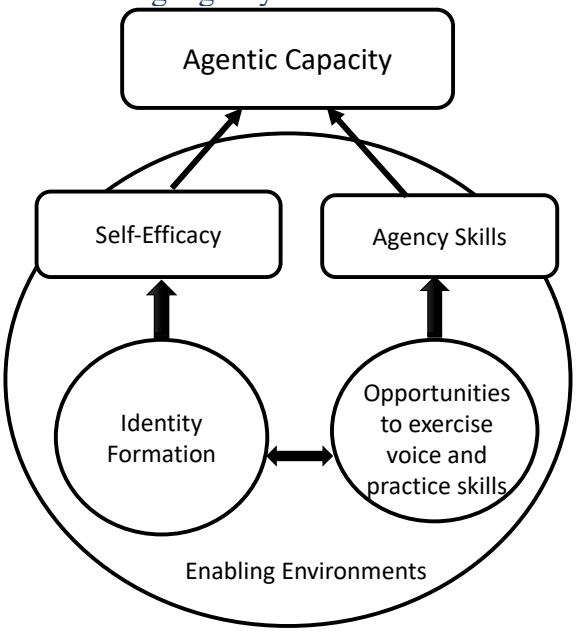
A Proposed Framework for Agentic Development

Given what we now understand about Agency as a combination of positive self-perceptions and concrete skills, as well as the relationship between identity formation and self-efficacy, it is not difficult to sketch a draft framework for the development of Agentic capacity in education practice (Figure 2). In the framework I postulate that programs that effectively cultivate Agentic capacity share two common components:

1. A theory of positive identity formation which leads to the development of self-efficacy; and
2. Opportunities for participants to learn and practice agency skills and assert their identities in relatively safe environments.

These two components reinforce each other. Turning to practice, we can see that the latter is often accomplished through life skills programs for young people that focus on skills building and the applied practice of social change. However, it is the former—positive identity formation—that is arguably the most critical ingredient and yet is much more difficult to accomplish. In the examples offered in this paper there is some evidence to suggest that positive identity formation happens naturally in environments of critical pedagogy and youth-directed or youth-driven learning. It is also possible that the development of self-efficacy can be an iterative process that happens in tandem with the acquisition and practice of agency skills, and vice versa. Regardless, it is important that practitioners have a clear working theory of how these two aspects are developed by their programs and test these theories through their monitoring and evaluation work.

Figure 2: Framework for Cultivating Agency in Life Skills Education Practice



Conclusion

A few things should be clear from this synthesis. First and foremost, Agency as defined in this paper offers a useful and locally adaptable framework for measuring current life skills programs. Understanding life skills from within an Agency frame suggests that focus on skills development should continue, but those skills should be identified according to the specific needs of each place based on the realities of culture, context and structural constraints.

Second, life skills programs should seek to foster positive self-perceptions, like self-efficacy and self-esteem. The evidence presented in this paper suggests that programs which actively facilitate positive identity formation and create space for the practice of self-expression and newly acquired skills are best positioned to cultivate the development of positive self-belief amongst adolescent girls. Further, alternative pedagogies are a particularly promising approach for doing the latter, because they focus on creating a safe power-neutral space where positive identity development and skills practice can take place in a risk free environment. This is a critical area for further research. While we know that alternative pedagogical practice can be an effective mechanism for fostering self-efficacy and Agency skills, more research is needed to fully catalogue and understand the mechanisms of change within the practices that do exist.

Third, Agency's unique relationship to identity formation makes adolescence a prime age for programs that build Agentic capacity. This is contrary to much of the current thinking—which identifies the "sweet spot" of many developmental processes as being in early childhood—underscoring the need for investment in research and programs for girls during older adolescence and during the secondary-school years.

Finally, this paper points to several areas of further research. Most immediately the need to understand in more detail the relationship between Agency and context. Which skills and self-beliefs are universally important for the development of Agentic capacity in girls from all socio-cultural and geographic contexts? Are there certain skills and self-beliefs that are more important for rural girls, for example than girls growing up in urban slums? How do country and culture influence the particular combination of skills and beliefs required for effective programming? Alongside contextually-specific understanding is the need to flesh out a more nuanced understanding of best practices for curricula and interventions that foster self-efficacy, both within the realm of critical pedagogy and without. What strategies have already been tested, and are there viable means for measuring success that are both culturally relevant and can quantify results and direct investment?

References

Advancing Girls' Education in Africa (AGE Africa). 2016. CHATS: *Creating Healthy Approaches to Success, Life Skills Curriculum.* Blantyre, Malawi.

Adely, Fida J. 2012. "Gendered Paradoxes: Educating Jordanian Women in Nation, Faith, and Progress." Chicago, IL: University of Chicago Press.

Acharya, Rajib, and Shveta Kalyanwala. 2009. "Broadening Girls' Horizons: Effects of a Life Skills Education Program in Rural Uttar Pradesh." New Delhi, India: Population Council.

Ahearn, L. M. (2010). Agency and language. *Society and Language Use*, *7*, 28–48.

Alvarado, G., Kapungu, C., Moss, C., Plaut, D., Reavley, N., & Skinner, M. (2017). *A Systematic Review of Positive Youth Development Programs in Low-and Middle-Income Countries*. YouthPower Learning and Making Cents International. https://pdf.usaid.gov/pdf_docs/PA00MR58.pdf

Austrian, K., Hewett, P. C., Soler-Hampejsek, E., Bozzani, F., Behrman, J. R., & Digitale, J. (2016). *Adolescent Girls Empowerment Programme: Research and evaluation mid-term technical report*. https://knowledgecommons.popcouncil.org/cgi/viewcontent.cgi?article=1560&context=departments_sbsr-pgy

Bajaj, M. (2015). 'Pedagogies of resistance' and critical peace education praxis. *Journal of Peace Education*, *12*(2), 154–166.

Bandura, Albert. 1982. "Self-Efficacy Mechanism in Human Agency." *American Psychologist* 37 (2): 122–47.

_____ 1989. "Human Agency in Social Cognitive Theory." *American Psychologist* 44 (9): 1175–84.

_____ 1993. "Perceived Self-Efficacy in Cognitive Development and Functioning." *Educational Psychologist* 28(2): 117–148.

_____ 2006. "Towards a Psychology of Human Agency." *Perspectives on Psychological Science* 1(2): 164–180.

_____ 2009. "Cultivate Self-Efficacy for Personal and Organizational Effectiveness--Indispensible Knowledge for Evidence-Based Management." In *Handbook of Principles of Organizational Behavior*, edited by Edwin A. Locke. West Sussex, UK: John Wiley & Sons Ltd.

Benson, P. L. (2003). Developmental assets and asset-building community: Conceptual and empirical foundations. In R. M. Lerner & P. L. Benson (Eds.), *Developmental Assets and Asset-Building Communities: Implications for Research, Policy and Practice* (pp. 19–43). Springer.

Bhalotra, Sonia, and Damian Clarke. 2013. Educational Attainment and Maternal Mortality (Background Paper). UNESCO. http://unesdoc.unesco.org/images/0022/002259/225945E.pdf

Borghans, Lex, Huub Meijers, and Bas Ter Weel. 2008. "The Role of Non-Cognitive Skills in Explaining Cognitive Test Scores." *Economic Inquiry* 46 (1): 2-12.

Brandtstadter, Jochen. 1999. “The Self in Action and Development: Cultural, Biosocial and Ontogenetic Basis of Intentional Self-Development.” In *Action and Self-Development: Theory and Research Through the Life Span*, edited by Jochen Brandtstadter and Richard M. Lerner, 37–65. Thousand Oaks, CA: Sage Publications, Inc.

Bronfenbrenner, U., & Morris, P. A. (1994). Ecological Models of Human Development. In *Readings on the development of children* (Vol. 1, pp. 37–43).

Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. “The Determinants of Earnings: A Behavioral Approach.” *Journal of Economic Literature* 39(4): 1137–1176.

Catalano, R. F., Skinner, M. L., Alvarado, G., Kapungu, C., Reavley, N., Patton, G. C., Jessee, C., Plaut, D., Moss, C., & Bennett, K. (2019). Positive youth development programs in low- and middle-income countries: A conceptual framework and systematic review of efficacy. *Journal of Adolescent Health*, *65*(1), 15–31.

Catalano, Richard F., M. Lisa Berglund, Jean A. M. Ryan, Heather S. Lonczak, and J. David Hawkins. 2004. “Positive Youth Development in the United States: Research Findings on Evaluations of Positive Youth Development Programs.” *Annals of the American Academy of Political and Social Science* 591: 98-124.

Cervantes-Soon, Claudia G. 2017. *Juárez Girls Rising: Transformative Education in Times of Dystopia.* Minneapolis, MN: University of Minnesota Press.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. “How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star.” *The Quarterly Journal of Economics* 126(4): 1593–1660.

Chisamya, Grace, Joan DeJaeghere, Nancy Kendall, and Marufa Aziz Khan. 2012. “Gender and Education for All: Progress and Problems in Achieving Gender Equity.” *International Journal of Educational Development* 32 (6): 743–55. https://doi.org/10.1016/j.ijedudev.2011.10.004.

Deyhle, Donna. 2009. *Reflections in Place: Connected Lives of Navajo Women.* Tuscon, AZ: University of Arizona Press.

Donald, Aletheia, Gayatri Koolwal, Jeannie Annan, Kathryn Falb, and Markus Goldstein. 2017. “Measuring Women’s Agency.” *Policy Research Working Paper*. World Bank Group Africa Region.

Dupuy, Kendra, Sosina Bezu, Are Knudsen, Sandra Halvorsen, Christina Kwauk, Amanda Braga, and Helyn Kim. 2018. “Life Skills in Non-Formal Contexts for Adolescent Girls in Developing Countries.” Center for Universal Education at the Brookings Institution.

Educate! 2014. *Socially Responsible Leadership Course* and *Teacher Training Curriculum.* Kampala, Uganda.

Educational Development Center. 2010. *Akazi Kanoze Youth Livelihoods Project.* Kigali, Rwanda.

Eloundou-Enyegue, P. M., & Stokes, C. S. (2004). Teen fertility and gender inequality in education: A contextual hypothesis. *Demographic Research*, *11*, 305–334.

Freire, Paulo. 1970. *Pedagogy of the Oppressed.* Translated by Myra Bergman Ramos. 30th Anniversary Edition. New York: Bloomsbury Academic.

Giroux, S., Eloundou-Enyegue, P., Sipple, J. W., & Tenikue, M. (2020). A Global Equalizer? Education and the Recent Economic Convergence of World Countries. *Annual Review of Comparative and International Education 2019*. https://books.google.com/books?hl=en&lr=&id=HNfpDwAAQBAJ&oi=fnd&pg=PT146&ots=nrzoE4wC4O&sig=I9uWJ7EXsTx2O-sIzrdr2wqxm_A#v=onepage&q&f=false

Giroux, Sarah, Fatou Jah, and Parfait Eloundou-Enyegue. 2010. "Globalization, Asymmetric Urbanization and Rural Disadvantage in Sub-Saharan Africa." In *Rural Education for the 21st Century: Identity, Place and Community*, edited by Kai A. Schafft and Alecia Youngblood Jackson. University Park, PA: Pennsylvania State University Press.

Giroux, H. A. (2001). *Theory and Resistance in Education: Towards a Pedagogy for the Opposition*. Bergin & Garvey.

Gupta, Neeru, & Mary Mahy. 2003. "Adolescent childbearing in sub-Saharan Africa: Can increased schooling alone raise ages at first birth?" *Demographic Research*, 8(4): 93-106.

Gutman, Leslie Morrison and Ingrid Schoon. 2013. "The Impact of Non-Cognitive Skills on Outcomes for Young People: Literature Review." London, UK: Institute of Education, University of London.

Haberland, Nicole. 2015. "The Case for Addressing Gender & Power." *International Perspectives on Sexual and Reproductive Health* 41(1): 31-42.

Hawkins, J.D. 1997. "Academic Performance and School Success: Sources and Consequences." In *Healthy Children 2010: Enhancing Children's Wellness*, edited by R.P. Weissberg, T.P. Gullotta, R.L. Hampton, B.A. Ryan, and G.R. Adams, 278–305. Thousand Oaks, CA: Sage Publications, Inc.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *NBER Working Paper Series* no. 12006. Cambridge, MA: National Bureau of Economic Research.

Holland, Dorothy, William Lachicotte, Jr., Debra Skinner, and Carole Cain. 1998. *Identity and Agency in Cultural Worlds.* Cambridge, MA: Harvard University Press.

hooks, bell. (1994). *Teaching to Transgress: Education as the Practice of Freedom*. Routledge, Taylor & Francis Group.

Hull, G., Jury, M., & Sahni, U. (2015). "Son Enough:" Developing Girls' Agency in India across Domestic and Civic Spheres through Feminist Media Practice. In B. Kirshner & E. Middaugh (Eds.), *# youthaction: Becoming Political in the Digital Age* (p. 81). Information Age Publishing.

Ibarrarán, Pablo, Laura Ripani, Bibiana Taboada, Juan Miguel Villa, and Brígida García. 2014. "Life skills, employability and training for disadvantaged youth: Evidence from a randomized evaluation design." *IZA Journal of Labor & Development*, 3(1): 10.

Jackson, C. Kirabo. 2012. "Non-Cognitive Ability, Test Scores, and Teacher Quality: Evidence from 9th Grade Teachers in North Carolina." *NBER Working Paper Series* no. 18624.

Jones, S., Schipper, Y., Ruto, S., & Rajani, R. (2014). Can your child read and count? Measuring learning outcomes in East Africa. *Journal of African Economies*, *23*(5), 643–672.

Kabeer, Naila. 1999. "Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment." *Development and Change*, 30(3): 435-464.

King, E. M., & Winthrop, R. (2015). *Today's Challenges for Girls' Education* (Working Paper Working Paper 90; Global Economy and Development at Brookings). The Brookings Institution.

Klasen, Stephan. 2002. "Low Schooling for Girls, Slower Growth for All? Cross-Country Evidence on the Effect of Gender Inequality in Education on Economic Development" in *The World Bank Economic Review*, 16(3): 345-373. The World Bank.

Kwauk, Christina, and Amanda Braga. 2017. "Translating Competencies to Empowered Action: A Framework for Linking Girls' Life Skills to Social Change." The Brookings Institution.

Larson, Reed W., and Rachel M. Angus. 2011. "Adolescents' Development of Life Skills for Agency in Youth Programs: Learning to Think Strategically." *Child Development*, 82(1): 277-294.

Learning Metrics Task Force. 2013. "Toward Universal Learning: Recommendations from the Learning Metrics Task Force Summary Report." *Learning Metrics Task Force*. Washington, DC: UNESCO & The Brookings Institution.

Lenton, Pamela. 2014. "Personality Characteristics, Educational Attainment and Wages: An Economic Analysis Using the British Cohort Study." *Sheffield Economic Research Paper Series*. Institute for Economic Analysis of Decision Making. https://www.sheffield.ac.uk/polopoly_fs/1.397279!/file/serps_2014011.pdf

Lerner, Richard M., Jacqueline V. Lerner, Jason B. Almerigi, Christina Theokas, Erin Phelps, Steinunn Gestsdottir, Sophie Naudeau, et al. 2005. "Positive Youth Development, Participation in Community Youth Development Programs, and Community Contributions of Fifth-Grade Adolescents: Findings From the First Wave Of the 4-H Study of Positive Youth Development." *The Journal of Early Adolescence*, 25(1): 17-71.

Lerner, Richard M., Christina Theokas, and Helena Jelicic. 2005. "Youth as Active Agents in Their Own Positive Development: A Developmental Systems Perspective." In *The adaptive self: Personal continuity and intentional self-development,* edited by W. Greve, K. Rothermund, & D. Wentura, 31-47. Ashland, OH: Hogrefe & Huber.

Leventhal, K. S., Gillham, J., DeMaria, L., Andrew, G., Peabody, J., & Leventhal, S. (2015). Building psychosocial assets and wellbeing among adolescent girls: A randomized controlled trial. *Journal of Adolescence*, *45*, 284–295.

Little, Todd D., C. R. Snyder, and Michael Wehmeyer. 2006. "The Agentic Self: On the Nature and Origins of Personal Agency Across the Lifespan." In *The Handbook of Personal Development*, edited by D. K. Mroczek & T. D. Little, 61-79. Mahwah, NJ: Lawrence Erlbaum Associates.

Malecki, Christine Kerres, and Stephen N. Elliot. 2002. "Children's Social Behaviors as Predictors of Academic Achievement: A Longitudinal Analysis." *School Psychology Quarterly* 17(1): 1-23.

Maslak, M. A. (Ed.). 2008. *The Structure and Agency of Women's Education*. Albany, NY: State University of New York Press.

Mensch, Barbara S., Barbara L. Ibrahim, Susan M. Lee, and Omaima El-Gibaly. 2003. "Gender-Role Attitudes among Egyptian Adolescents." *Studies in Family Planning* 34(1): 8-18.

Mirón, Louis F., and Mickey Lauria. 1998. "Student Voice as Agency: Resistance and Accommodation in Inner-City Schools." *Anthropology & Education Quarterly*, 29(2): 189-213.

Murphy-Graham, E. and Lloyd, C. 2016. "Empowering adolescent girls in developing countries: The potential role of education." *Policy Futures in Education,* 14 (5): 556 - 577.

Murphy-Graham, Erin, and Graciela Leal. 2015. "Child Marriage, Agency, and Schooling in Rural Honduras." *Comparative Education Review* 59(1): 24-49.

Osborne Groves, M. 2005. "How important is your personality? Labor market returns to personality for women in the US and UK." *Journal of Economic Psychology*, 26: 827-841.

Population Council. 2013. *Health and Life Skills Curriculum.* Lusaka, Zambia.

Psaki, Stephanie R., Barbara S. Mensch, and Erica Soler-Hampejsek. 2017. "Associations Between Violence in School and at Home and Education Outcomes in Rural Malawi." *Comparative Education Review,* 61(2): 354-90.

Sahni, Urvashi. 2017. *Reaching for the Sky: Empowering Girls Through Education.* Washington, DC: Brookings Institution Press.

Santhya, K.G., A.J. Francis Zavier, and Shireen J. Jejeebhoy. 2015. "School Quality and Its Association with Agency and Academic Achievements in Girls and Boys in Secondary Schools: Evidence from Bihar, India." *International Journal of Educational Development* 41(Supplement C): 35-46.

Search Institute. 1997. "40 Developmental Assets for Adolescents (ages 12-18)." Minneapolis, MN. Accessed at www.search-institute.org on April 4th 2018.

Shah, P. P. (2011). Girls' Education and Discursive Spaces for Empowerment: Perspectives from Rural India. *Research in Comparative and International Education*, *6*(1), 90–106. https://doi.org/10.2304/rcie.2011.6.1.90

Sperling, Gene B., and Rebecca Winthrop. 2016. *What Works in Girls' Education: Evidence for the World's Best Investment.* Washington, DC: Brookings Institution Press.

Stevens, Robert J., and Robert E. Slavin. 1995. "The Cooperative Elementary School: Effects on Students' Achievement, Attitudes, and Social Relations." *American Educational Research Journal* 32(2): 321-51.

Temin, M., & Heck, C. J. (2020). Close to Home: Evidence on the Impact of Community-Based Girl Groups. *Global Health: Science and Practice*, *8*(2), 300. https://doi.org/10.9745/GHSP-D-20-00015

The Academy for Educational Development. 2008. *Ambassador Girls' Scholarship Program Guide to Mentoring and Life Skills.*

Unterhalter, E., Arnot, M., Lloyd, C., Moletsane, L., Murphy-Graham, E., North, A., Parkes, J., & Saito, M. (2014). *Interventions to Enhance Girls' Education and Gender Equality: Education Rigorous Literature Review*. Department for International Development London.

Santen, José C. M. 2010. "'My "veil" does not go with my jeans': veiling, fundamentalism, education and women's agency in northern Cameroon." *Africa*, 80(2): 275-300.

Warner, Ann, Kirsten Stoebenau, and Allison M. Glinski. 2014. "More Power to Her: How Empowering Girls Can Help to End Child Marriage." Washington, DC: International Center for Research on Women.

Welsh, Mara, Ross D. Parke, Keith Widaman, and Robin O'Neil. 2001. "Linkages Between Children's Social and Academic Competence: A Longitudinal Analysis." *Journal of School Psychology* 39 (6): 463-81.

Winthrop, Rebecca, and Eileen McGivney. 2016. "Skills for a Changing World: Advancing Quality Learning for Vibrant Societies." Washington, DC: The Brookings Institution, Center for Universal Education.

Wodon, Quentin T., Chata Male, Kolobadia Ada Nayihouba, Adenike Opeoluwa Onagoruwa, Aboudrahyme Savadago, Ali Yedan, Jeff Edmeades, Aslihan Kes, Neetu John, Lydia Murithi, Mara Steinhaus, and Suzanne Petroni. 2017. "Economic Impacts of Child Marriage: Global Synthesis Report." World Bank and International Center for Research on Women.

Wodon, Q., Montenegro, C. E., Nguyen, H., & Onagoruwa, A. (2018). *Missed Opportunities: The High Cost of Not Educating Girls* (Girls Notes Series). World Bank.

Zins, Joseph E., Roger P. Weissberg, Margaret C. Wang, and Herbert J. Walberg (Eds.). 2004. *Building Academic Success on Social and Emotional Learning: What Does the Research Say?* New York NY: Teachers College Press.

PAPER #2: ASSESSMENT OF A PRACTITIONER-DERIVED FRAMEWORK FOR MEASURING GIRLS' AGENCY IN EAST AFRICA

Submitted for 2nd review Jan 18th 2022 to the *Comparative Education Review,* in accordance with a 'revise and resubmit' decision received in October 2021.

Aubryn Allyn Sidle and Brenda Oulo

Abstract:

Although soft skills are considered an important part of quality education, there is little consensus on which skills matter for which populations, and few cross-culturally valid measures of such skills. We propose a conceptual and methodological remedy to this quandary through an investigation of "agency" as an outcome measure for adolescent girls' life skills programs. We engage community-based practitioners in East Africa who identify agency to be a priority outcome for girls and derive a four-dimensional structure of agency from theory, practitioner expertise, and empirical evidence. We develop and pilot a regional measure of agency for adolescent girls, utilizing confirmatory factor analysis to confirm the underlying structure of our instrument and to investigate its use as a pre-post measure. Analysis was conducted using data collected from 1,953 girls in four countries. Results indicate a four-latent factor structure representing the construct of agency. The implications of our study suggest that researchers should prioritize investigating the regional limits to validity of contextually specific soft skills measures over universal metrics, and that measures should be developed in collaboration with local communities.

Acknowledgements: We would like to acknowledge the individual organizations and members of the AMPLIFY Girls (AMPLIFY) network who collaborated on this research: AkiraChix, Chalbi Scholars, Dandelion Africa, Elohim Development Association, Gashora Girls Academy, Girls to Lead Africa, Girls Livelihood and Mentoring Initiative, Girl Up Initiative Uganda, Jifundishe, Komera, Malkia Foundation, Riley Orton Foundation, Sega Girls Academy, The Action Foundation of Kenya, The Girls Foundation of Tanzania, The Streets Ahead Children's Center Association, and WISER Girls. We would also like to thank our regional research liaisons: Dr. Eugenia Kafanabo at the University of Dar es Salaam (Tanzania), Dr. John Bosco Isunju at Makerere University (Uganda), Dr. Wenceslas Nzabalirwa at the University of Rwanda, Dr. Alfred Kitawi at Strathmore University (Kenya), and AMPLIFY's Research Assistant, Ninjel Osano. This research would not have been possible without their support. We would also like to thank Echidna Giving for their generous support.

1. Introduction: Soft Skills, Non-Cognitive Skills or Social Emotional Learning for Girls Education?

Early discussions on the framing of Sustainable Development Goal 4 (SDG4) and the "quality learning for all" agenda identified a broad range of "soft skills" competencies as being important for quality education (Learning Metrics Task Force, 2013). Yet few of these skills were formally adopted onto the final SDG4 agenda, even though soft skills are necessary for the achievement of formal SDG4 targets such as "promoting global citizenship and sustainability" (Care & Anderson, 2016). Whereas numeracy and literacy are taught and measured by school systems around the world, soft skills are notoriously hard to assess, in part because of the practical difficulty of creating contextually specific definitions and associated test-scales for qualitative capacities such as "empowerment" or "resilience."

Soft skills are also not a monolith, having been defined in overlapping but divergent ways across disciplines. In human capital debates, for example, soft skills are competencies, behaviors, attitudes, and attributes that promote critical thinking, collaboration, work ethic, goal achievement, and other capacities that matter for performance on the globalized labor market (Lippman et al., 2015; Burnett and Jayaram, 2012). Related to soft skills, labor economist James Heckman popularized the term "non-cognitive" skills to refer to the broad set of non-academic competencies that are causally linked to improved labor market performance (Heckman et al., 2006; Kautz & Heckman, 2014). The corollary set of skills in education are called social emotional learning (SEL) competencies and refer to the skills associated with young peoples' positive self-concept and the social interactions that support learning (Zins et al., 2004). Recent efforts have classified SEL competencies into six domains: cognitive, emotional, social, values, perspectives, and identity (Brush et al., 2022). Arguably the distinction between soft skills, SEL, and non-cognitive skills is

one of emphasis, related to the specific field that first conceptualized the term and the associated outcome of interest.

In the international girls' education sector however, soft skills and SEL initiatives are frequently combined in the form of life skills education programs that target outcomes across sectors including education, economic empowerment, and health (Nasheeda et al., 2019; Brush et al., 2022). In this paper, we offer a useful measure for evaluating informal life skills education programs and use the term "soft skills" in the broadest sense, to mean the range of SEL and non-cognitive competencies that contribute to improving girls' well-being across these diverse outcome areas. We address the problems of both consensus and measurement in girls' education by collaborating with practitioners of the AMPLIFY Girls collective in East Africa, who identify "agency" as a key outcome of interest for life skills programs. With these practitioners, we produce a robust conceptualization of agency as a context-dependent construct and produce a regional measure of agency for evaluating girls' life skills programs in East Africa.

2. The Importance of Agency for Adolescent Girls and Shortcomings of Existing Measures

Randomized controlled trials of life skills interventions in low income countries have shown that soft skills programming has a positive effect on girls' job skills and labor market outcomes,[26] improves sexual and reproductive health (Temin & Heck, 2020; Erulkar et al., 2013), and makes a positive contribution to psychosocial well-being (Temin & Heck, 2020; Leventhal et al., 2015). Yet several gaps remain in this growing literature. First, with a few notable exceptions, soft skills are largely studied as a nebulous conglomerate of traits and attributes with wide variation in domain and with few attempts to link specific skills to specific outcomes (Jackson, 2012).

[26] (Catalano et al. 2019 ; Ellis and Chafin 2015; Ibarrarán et al. 2014; Bandiera et al. 2014; Dickson and Bangpan 2012)

Further, although there have been recent attempts to define the dimensions of soft skills for SEL (Brush et al., 2022), few have offered a robust conceptualization of what skills or frameworks matter for vulnerable populations as opposed to youth more broadly.

Agency (defined as the ability to act) is a soft skill construct with important implications for vulnerable populations such as women and adolescent girls. Naila Kabeer's work (1999) has widely popularized the notion of agency as a prerequisite for women's economic and social empowerment and, as such, agency is frequently viewed by researchers and practitioners as a means towards an end, rather than as an outcome in itself. Thus, measures of women's agency have been relatively one-dimensional and focused on singular indicators that might show if a woman *has* agency, rather than trying to understand and appropriately capture the dimensions of agency as a construct. These measures commonly focus on assessing the ability of women to make autonomous decisions in the context of marriage and household (Donald et al., 2020).

Evidence from youth development shows that agency is not only important for women however, but also for adolescent girls (Banati et al., 2021; Catalano et al., 2019), and that "agency" should be the expected outcome of empowering and high-quality education (Murphy-Graham & Lloyd, 2016). Yet most existing measures—such as the Women's Empowerment Agricultural Index, the DHS' SWPER, or the 5-question agency index (Jayachandran et al., 2021)—assess agency within the context of a specific domain of adult life and livelihood (i.e. agriculture, or the marital household) and are thus inappropriate for adolescents whose domains of influence are different from their married counterparts. Where decision-making autonomy is assessed in adolescent girls, we see measures that proxy for a girl's ability to negotiate around marriage and school attendance, but which exclude aspects of her well-being beyond sexual and reproductive health.

Measures of agency in any sphere which only assess autonomous decision-making also ignore the necessary psychological dimensions of the construct. Self-perceptions about a person's ability to act are an integral part of agency (Bandura, 2006) and must be included in robust measures of the concept (Donald et al., 2020). A 2017 World Bank review of existing measures of women's agency found deficiencies in the conceptualization and content description of agency in current measures, noting that they typically conceived agency as a one-dimensional proxy for "choice" and were untested/unvalidated for use in specific socio-cultural environments (Donald et al., 2017, pp. 2–3). In the same year, a USAID report reviewed nearly 250 tools measuring soft skills in international youth development programs, concluding that:

> … to improve the state of youth soft skill measurement, a soft skill assessment [should] be developed that draws from the universe of existing tools, is designed specifically for program use … [It] should be short and easy to administer, translated into languages needed for programs in Latin America, Africa, Middle East, and Asia … [and] developed and pilot tested in multiple international program contexts … [It] might incorporate multiple methods to mitigate the shortcomings of self-report … [and] implementers should be included from the outset of tool development. (Galloway et al., 2017, p. 56)

We tackle the problems of conceptual clarity, purpose of use, rigor of testing, and validation and practitioner inclusion through a collaborative research partnership with AMPLIFY Girls—a network of eighteen community-based program implementers in East Africa. Members of the collective operate informal life skills education programs for girls in both community and school-based settings in Kenya, Tanzania, Rwanda, and Uganda.

3. Agency as a Multidimensional and Context-Dependent Construct

In 2018, AMPLIFY member organizations launched an effort to identify shared metrics that would enable them to track their collective impact on adolescent girls across the region. Specifically, AMPLIFY wanted to create a quantitative and easily reportable measure for evaluating girls' soft skills outcomes, because they felt these skills were some of the most important factors for improving girls' well-being in the long-term. Utilizing Group Concept Mapping (Trochim & Kane, 2005) as the methodological tool for conceptualizing shared metrics, the final output showed that AMPLIFY organizations unanimously identified girls' agency as the most important outcome for measuring the impact of life skills education programs. Furthermore, AMPLIFY's conceptualization of the agency domain was made up of thirty-six statements representing both skills and beliefs related to the capacity to make change, suggesting that practitioners' view of agency was multi-dimensional (AMPLIFY, 2019).[27]

The human development literature agrees with a multi-dimensional view of agency. Psychologist Albert Bandura developed a comprehensive theory of human agency, defining it as two basic functions that work in tandem: a set of self-perceptions or self-beliefs and a set of coordinating skills (Bandura, 2006). Development practitioners have long relied on Kabeer's definition of agency as the "power-to" (Kabeer, 1999, p. 498) which she argues is made up of two dimensions—a woman's ability to define her own goals and to make her own life choices—and, more recently, Donald et al. (2017, p. 6) who expand Kabeer's definition to understand agency as aspirational goal setting combined with self-perceived ability and action towards goal achievement. Implicit in both conceptualizations of agency is an acknowledgment that underlying

[27] A detailed description of the Group Concept Mapping study which led to this view of agency as a multi-dimensional outcome for AMPLIFY can be found in the cited report.

"the ability to act" is a complex set of skills and beliefs that both influence and are influenced by a person's environment.

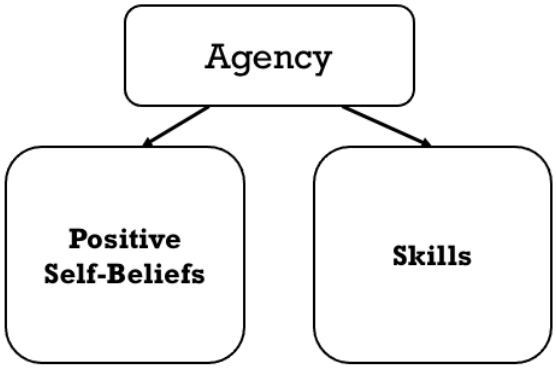
Specifically, Kabeer argues that agency requires decision-making "even in the face of opposition from others" (1999, p. 498), suggesting that agency is a flexible construct, the specific make-up of which is dependent on the types of oppositional forces that dominate a person's life. For adolescent girls in East Africa, oppositional force importantly includes the intersection of poverty and gendered norms and expectations. Access to education, for example, has been well documented as a function of household income and per capita gross domestic product (Lloyd & Hewett, 2009). In East Africa, household poverty is also the primary predictor for the highest rates of teen pregnancy and early marriage for girls (Wado et al., 2019; Neal et al., 2015). Such facts indicate that girls from resource-limited backgrounds face gendered challenges in addition to resource-based challenges in finishing school.

Indeed, in each of our study countries, the poorest quintile of girls fail to transition to secondary school at a rate that is almost 2:1 compared to boys in the same demographic (UNESCO, 2021), indicating that a girl's agency is importantly limited by this intersection of her experiences with poverty and gender. In a study of young women's education programs in Honduras, Murphy-Graham and Leal (2015) found that girls exercise their agency based on a complex set of sociocultural conditions and constraints, concluding that effective education interventions must attend to these aspects of girls' context in order to be successful. Thus we argue that girls' agency is "mediated" by the socio-cultural forces in the environment, especially gendered norms and expectations (Ahearn, 2010), in addition to the socio-economic forces which structure her opportunities. From this view, a girl's agency is context-dependent and should be defined by the

specific skills and beliefs that help her effectively navigate and mitigate these forces in her particular setting.

Applying this "dependent" view of agency, we take Bandura's basic framework of agency as beliefs plus skills to be essentially correct, but we posit that the specific skills and beliefs that make up agency for girls are mediated by aspects of macro socio-economic structure, rigidity of gender norms, and girls' beliefs about those norms. Our survey development process is informed by this conceptualization of agency, as captured by Figure 1, where agency is presented as a two-dimensional construct made up of self-beliefs and skills, the domains of which are context-dependent (Sidle, 2019).

Figure 1 *Simple Conceptual Framework for Agency*

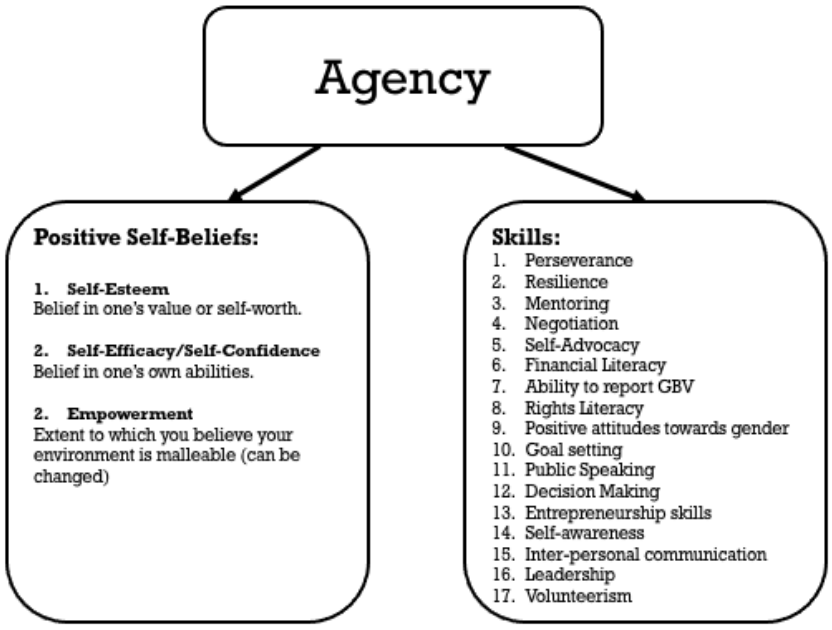


Adapting our Simple Conceptual Framework of Agency to AMPLIFY's Setting

In order to adapt this simple conceptual framework to AMPLIFY's setting, we asked practitioners representing each of the eighteen AMPLIFY organizations to identify and rank the specific skills and self-beliefs which make up agency for the adolescent girls they serve. For more ambiguous constructs such as empowerment and self-confidence, practitioners defined these skills or beliefs in their own words. Definitions were reviewed and synthesized by the researchers into distinct categories, yielding a complete list of twenty skills and four beliefs which AMPLIFY

practitioners identified as important for agency outcomes. Seventeen of these skills were selected by a majority of members (60%), and thus were included as the basis of a draft agency survey alongside three self-beliefs: self-esteem, empowerment (defined by AMPLIFY members as "the extent to which an individual believes they have control over or can change their environment"), and self-efficacy. The fourth belief, "self-confidence," was defined by AMPLIFY members identically to self-efficacy, and so was dropped. Our simple two-dimensional conceptual framework of agency with the specific skills and beliefs making up agency according to AMPLIFY, is detailed in Figure 2.

Figure 2 *AMPLIFY Two-Dimensional Conceptual Framework of Agency*



Survey Development

With the two-dimensional conceptual framework as the starting point, we drafted the survey by selecting pre-existing psychometric scales of the seventeen skills and three beliefs illustrated in Figure 2. Emphasis was placed on selecting measures that had been previously validated for use in countries outside of the Global North and/or with non-European or non-white populations. Measures meeting these criteria included: The Rosenberg Self-Esteem Scale

(Rosenberg, 1965), the general self-efficacy scale (Schwarzer & Jerusalem, 1995), the empowerment scale (Torre, 1986), goal setting, interpersonal communication, and decision-making scales developed by Mincemoyer and Perkins (2005), the Gender Equitable-Men's Scale (C-Change, 2011), Duckworth and Quinn's Grit scale measuring "perseverance of effort" (2009), and the self-awareness portion of Scheier and Carver's self-consciousness scale (Scheier & Carver, 1985; Scholz et al., 2002). Several skills did not have pre-existing psychometric measures or measures that matched practitioners' definitions of the skills. In these cases, researchers drew from AMPLIFY members' existing surveys and literature.

In a workshop conducted in April 2019, AMPLIFY practitioners reviewed the resulting draft survey for the integrity of the questions vis-à-vis the definitions of skills or beliefs developed by AMPLIFY practitioners and for accessible and regionally appropriate English. AMPLIFY practitioners serving as content reviewers during this workshop were viewed as the programmatic experts of their respective organizations and were living and working in the communities where their programs operate. The vast majority (80%) were also from the communities of practice. As a result of these discussions, two skills that had not been included in the original fifteen (problem solving and conflict resolution) were returned to the survey by practitioners in workshops, and three other skills (entrepreneurship, financial literacy, and mentoring) were dropped because they were skills deemed important for vocations that *could* support an individual's agency but were not necessarily *essential* to agentic capacity. Resilience, self-advocacy, and negotiation were dropped from the survey because two of these skills (resilience and negotiation) were seen as overlapping substantially with other constructs on the survey, while self-advocacy was identified as a proxy for already having agency rather than as an actual indicator of agency. Three skills (gender rights,

positive gender attitudes, and the ability to recognize gender-based violence) were combined into one belief: gender beliefs.

The draft survey was tested with 350 respondents in Tanzania and Uganda, and the resulting data was used to examine the response variability of each item. Items with little to no variability were dropped from the survey, along with items that, according to AMPLIFY members, did not match the given definition of the attribute it measures. This resulted in a reduction of the survey from 101 items to 77. Table 1 shows the overview of items and constructs on the survey and their evolution from draft to pilot to the final parsimonious version presented after the analysis described in this paper.

Table 1 *Overview of Items on Each Survey Version by Construct*

	Draft Survey	**Pilot Survey**	**Final Parsimonious Survey**
Beliefs	**34**	**24**	**18**
self-efficacy	10	8	6
self-esteem	11	7	3
empowerment	13	9	3
gender beliefs	not included	not included	6
Skills	**67**	**53**	**22**
gender attitudes, rights and GBV awareness	17	11	moved
leadership Skills	6	5	5
interpersonal communication	6	5	1
volunteerism	5	4	3
public speaking	3	3	2
conflict resolution	not included	5	2
goal setting	4	3	1
problem solving	not included	3	2
decision making	4	4	1
perseverance	5	4	1
self-awareness	6	6	4
resilience	3	dropped	dropped
self-advocacy	2	dropped	dropped

Negotiation	3	dropped	dropped
entrepreneurship & financial literacy	3	dropped	dropped
Total Items	101	77	40

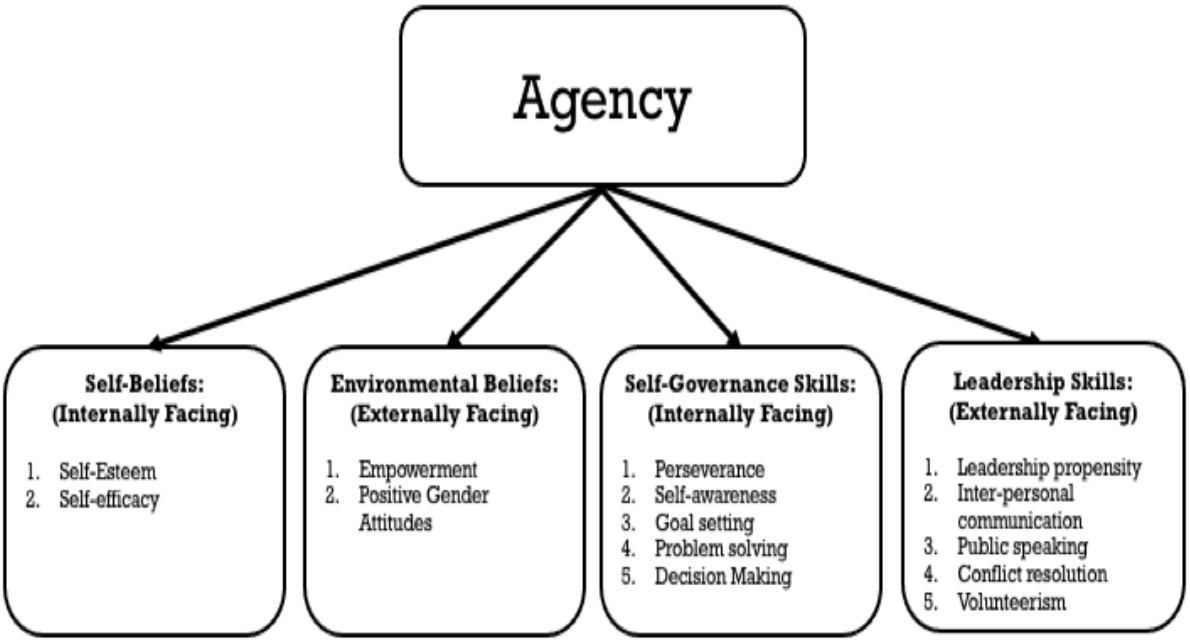
Developing a Four-Dimensional Conceptual Framework for Agency

To conceptualize the dimensions of our instrument more rigorously, we asked AMPLIFY practitioners to produce their own conceptual framework of agency. In four small focus group discussions comprised of the program leadership of each organization, AMPLIFY members grouped the thirteen attributes which they had identified as making up "agency" into dimensions that they felt represented the concept. These discussions produced a consensus across groups that the skills and beliefs making up agency should be further specified into those that are internally and externally facing. In other words, AMPLIFY members felt that the skills that were important for influencing the environment (externally-facing) were distinct from the skills necessary for organizing and coordinating a person's own action (internally-facing). Practitioners had a similar view of beliefs and organized them as beliefs about the self (internally-facing) and beliefs about the environment and its malleability (externally-facing). AMPLIFY members titled these categories as leadership skills, self-governance skills, self-beliefs, and environmental beliefs.

The resulting framework shown in Figure 3 depicts the result of this qualitative process, which refined our conceptual framework of agency from two dimensions to four. We argue this four-dimensional conceptual framework is a more accurate and contextualized view of the underlying structure of agency for East African girls living in AMPLIFY's communities. The framework essentially argues that, in order to have agency, girls need both to believe in themselves and their own abilities (self-beliefs), believe that their environments are malleable (environmental beliefs), be able to organize their abilities and aspirations (self-governance skills), and have the skills to influence others (leadership skills). We use this theory-driven conceptual framework as

our hypothesized model of the underlying structure of agency. To confirm this structure, we compare the theory-driven model to both the simple framework depicted in Figure 3 and a data-driven structure derived from exploratory factor analysis, which is described in the next section.

Figure 3 *AMPLIFY Four-Dimensional Conceptual Framework of Agency*



4. Methodological Approach

Our approach to understanding the underlying structure of agency relies on the "argument based approach" to validity detailed by Kane (2013). In this paper, we begin to build a case for the validity of our instruments' interpretation and intended use (Sireci, 2014). We establish evidence related to the content of our instrument in the form of theory, literature, and the participatory processes already discussed. We seek to confirm the structure of agency as a construct through confirmatory factor analysis, comparing our hypothesized theory-driven structure (Figure 3) to two others: one that is primarily literature derived (Figure 1), and one that is empirically derived using exploratory factor analysis (EFA). Further, we assess the validity of our instrument's intended use as a pre-post outcome evaluation measure by determining if configural, scalar, and metric invariance is supported in pre-post groups of participants.

Participants

Our data is derived from AMPLIFY's beneficiary pool. The final agency survey was administered to adolescent girls attending one of fifteen of the eighteen AMPLIFY member organizations' programs between August 2019 and February 2020. Surveys were translated and administered in three languages (English, Kiswahili, and Kinyarwanda) based on participants' language comprehension. All survey versions were professionally translated and back translated to ensure that the original meaning of the question was preserved during translation. Professional translations were also cross-checked by local language speakers on staff with AMPLIFY organizations. Although the majority of Rwandan and Tanzanian participants took the survey in their local languages, nearly all Kenyan and Ugandan respondents (and a majority of participants––overall 65%) took the survey in English. The data is cross-sectional, representing either students just entering AMPLIFY members' programs (baseline) or those who were in their final few months/weeks of members' programs (endline).

Materials

The AMPLIFY agency survey is a 77-item self-report questionnaire with response options given as five Likert scale ratings asking respondents to answer based on either frequency or on their level of agreement with a question.

Procedure

The researchers obtained ethical approval and research permission from local ethics committees and research-permitting bodies in Kenya, Uganda, Rwanda, and Tanzania. Informed consent was obtained from all participants before survey administration. The survey was administered to eligible participants in Kenya, Uganda, Rwanda and Tanzania. For the analysis, the dataset was stratified by survey type (baseline versus endline) and by country, and then randomly split into 1/3 (651 observations) and 2/3 (1,302 observations). Exploratory and principal

factor analysis was conducted on the larger sub-sample, and confirmatory factory analysis using structural equation modelling was conducted on the sample N=651.

5. Statistical Analyses

Analysis was performed using Stata version 16 (StataCorp) and MPLUS version 8.7 (Muthén & Muthén, 2017) for structural equation modelling (SEM) modelling.

Establishing the dimensions of agency

Current best practice in survey development and validation suggests that confirmatory factor analysis (CFA) is an effective way to hypothesize and confirm or to reject the underlying structure of relationships between constructs (Lopez et al., 2015; DeVellis, 2016). This method relies heavily on the robust conceptualization of the framework to be tested. For our analysis, we rely on the theory-driven conceptual framework developed in collaboration with AMPLIFY practitioners and depicted in Figure 3. We impose the four-factor theory-driven structure on the data to perform CFA using diagonally weighted least squares estimator (WLSMV in Mplus), designed for SEM model fitting with categorical data (Muthén & Muthén, 2017) and model error terms as independent.

To achieve parsimony, we use factor loadings and assessment of specific factor indicators to refine the scale by dropping items that did not load at .5 or greater onto the four factors but by retaining five items with loading between 0.4 and .49 that were crucial to the definition of the construct (Costello & Osborne, 2005). The final agency survey we present is a parsimonious 40-item questionnaire that represents the four dimensions of agency as a construct made of up of four dimensions: self-beliefs, environmental beliefs, self-governance skills and leadership skills. The final number of items on the parsimonious survey can be viewed in Table 1.

To produce an alternative model for comparison, we perform an EFA on the larger dataset (N=1,301) to establish the latent factor structure of our instrument based exclusively on

commonalities within the data. Factor extraction was accomplished using Kaiser's criterion, which retains factors with an eigen value above 1 after an examination of the scree plot. Further, we performed oblique promax rotations to allow for correlations between factors and extracted a four-factor solution that accounted for 64% or more of the variability. Items that did not load at .4 or greater, or cross-loaded at above .40, were removed.

Model Comparison

We compare model 3 (data-driven), resulting from the EFA derived factor structure described above; model 2 (theory-driven), corresponding to the imposed four-dimensions of agency as developed by AMPLIFY practitioners; and model 1, a simpler and uncorrelated two-factor model broadly representing skills and beliefs that approximates the simple conceptual framework (Figure 2). To determine which model best fits the data, the following fit statistics were obtained and compared: the chi-square with degrees of freedom, the root-mean-square error of approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), comparative fit index (CFI), and Tucker-Lewis Index (TLI).

When comparing exact fit statistics, a lower chi-square value indicates a better fit, given an equal number of degrees of freedom, although an insignificant chi-square suggests there is no difference between the model-implied structure and the data. We compare and adjudge approximate fit statistics using the conventional cutoffs for RMSEA, SRMR, CFI, and TLI statistics, but we acknowledge that questions have been raised about their use with the diagonally weighted least squares estimators often used to model categorical data (Xia & Yang, 2019). However, because no better conventions have been proposed and because we evaluate model fit comparatively rather than a priori, we are confident that the widely used conventions are also useful for our setting.

Accordingly, convention suggests that approximate-fit statistics should be evaluated based on both conservative and liberal cutoff values because fit indices may be affected by factors such as sample size and data distributions (Schermelleh-Engel et al., 2003). In our case we consider RMSEA of 0.05 or less representing a close fit of the model with values of up to 0.08 viewed as acceptable (Schermelleh-Engel et al., 2003). Similarly we evaluate the SRMR values based on the convention that 0.8 or less is considered a good fit (Hu & Bentler, 1999), and evaluate CFI and TLI statistics based on the liberal cutoff of .90 and conservative cut-offs of CFI/TLI ≥ 0.95 (Schermelleh-Engel et al., 2003).

Next, we examine measurement invariance across pre/post participant groups by performing multigroup CFA to assess whether measurement invariance is supported. We firstly examine configural, metric, and scalar invariance of the first order latent factor model, then proceed to establish invariance of the 2nd order factor model. We compare model chi square statistics using the difftest option in Mplus and compare CFI values for the configural, scalar, and metric models. The results are reported in Table 5.

6. Results

Data

A description of the respondents' demographics can be found in Table 2, disaggregated by country. Overall, there were 1,892 respondents with ages ranging from 10 to 25 years, with a mean age of 16.5 years (SD 2.8). Data was collected from four countries, with the majority of respondents from Kenya (37%), and the least from Uganda (13%). Sixty-seven percent of the respondents were from rural communities (1,264). Sixty-two percent represented students just entering programs designated as pre-intervention, with the remainder just completing programs (post-intervention). On average, respondents came from households with relatively low rates of education: 41% reported having parents who had only finished primary school, or who had never

attended school at all. Sixty-five percent took the survey in English, 23% in Kiswahili, and 12% in Kinyarwanda. For most of the participating organizations, all students entering or exiting their life skills programs were included in the sample, as the total number of participants was less than 100. For the remainder of organizations, a simple random sampling technique was employed to select a representative sample of an organization's beneficiaries to take the survey.

Table 2 *Summary of Data*

Demographic Characteristics	**Kenya**	**Tanzania**	**Rwanda**	**Uganda**	**Total**
N	775	572	363	243	1,953
Age (range)	10-24	11-25	12-25	10-24	10-25
Age (mean)	15.4	16.1	18.2	15.1	16.2
Urban/Peri-urban	147	13	158	156	474
Rural	628	559	205	87	1,479
parents finished primary school or less	308	289	127	77	801
parents finished secondary	260	127	21	103	511
parents attended higher education	116	77	50	26	269
% parents education missing	12%	14%	45%	15%	19%
Survey Characteristics	**Kenya**	**Tanzania**	**Rwanda**	**Uganda**	**Total**
Baseline	403	347	284	195	1,229
Endline	372	225	79	48	724
English	766	141	111	243	1,261
Swahili	9	431	0	0	440
Kinyarwanda	0	0	164	0	164

Confirming the Factorial Structure of Agency

Model 1, representing the simple conceptual framework of agency (Figure 1), did not converge. Both the theory-driven model (model 2) and factor analysis from the data-driven model 3 suggests that agency is a four-dimensional measure. In the data-driven structure, oblique promax rotation produced four factors most consistent and interpretable that account for 64% of the total variance. Results of EFA saw an outcome of a reduction of forty items that did not produce factor loadings of greater than .4. Both the scree plot and the factor correlation matrix of the data-driven

model can be found in the Appendix as Figure 4 and Table 3 respectively. Factor correlation coefficients ranged from .1 to .41, suggesting weak correlation between all factors in the data-driven model.

Table 4 illustrates absolute and relative fit statistics for model 3 (data-driven) and model 2 (theory-driven). We present the final preferred model representing the structure of agency following the convention that *both* absolute and relative fit statistics are important for adjudging model fit (Schermelleh-Engel et al., 2003; Mulaik, 2007). On the basis of absolute fit statistics, the theory-driven model (2) provides a better fit of the data with RMSEA of .04 (compared to .05 for model 3) and an acceptable SRMR value of 0.062. The relative fit statistics are also better for model 2 compared to model 3, with higher CFI value of .917 compared to .0909 and TLI 0.912 versus 0.901.

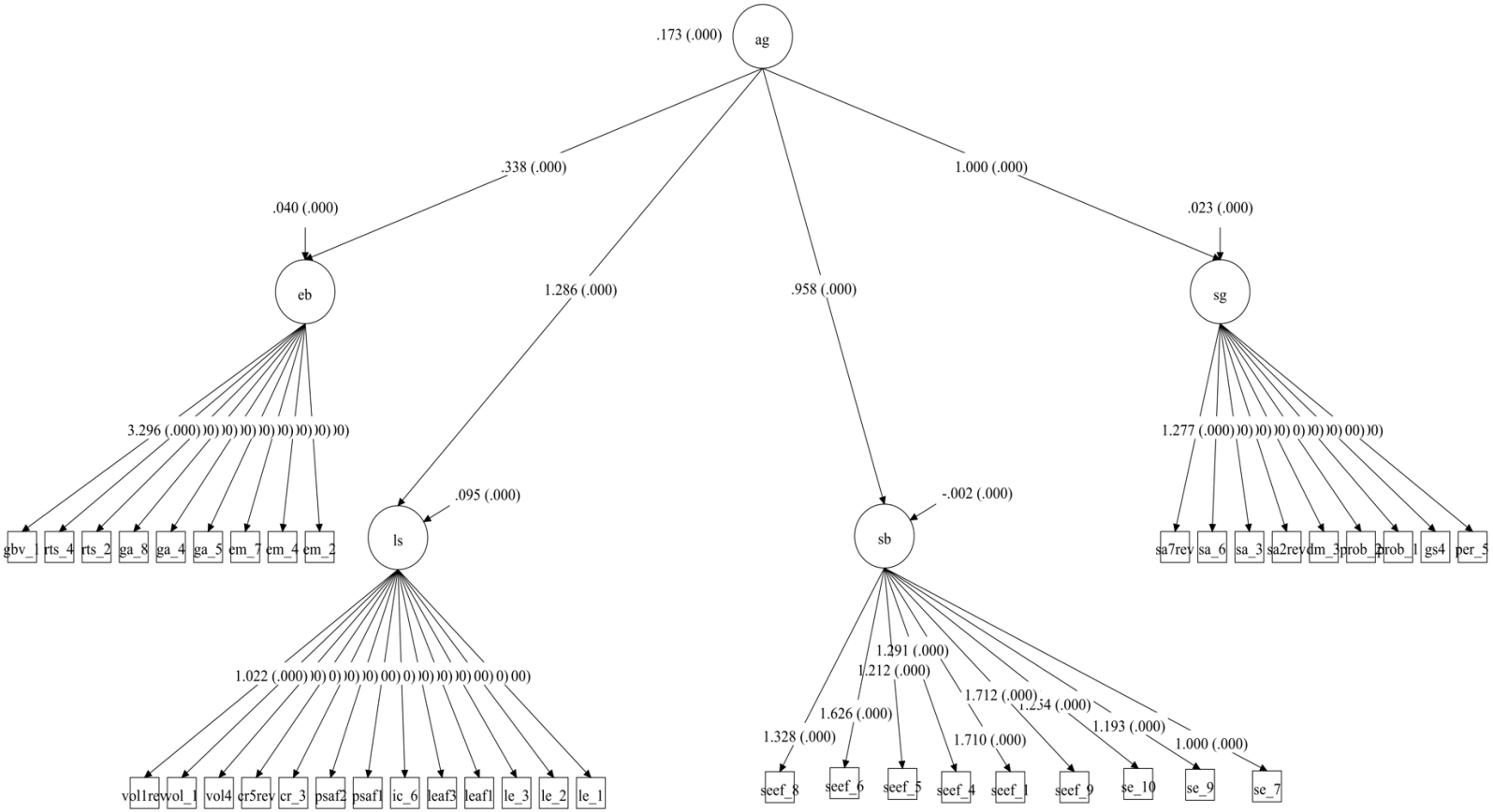
Table 4 *Exact and Approximate Fit Statistics for Model 2 and Model 3*

	Chi-Square	**Degrees of freedom**	**RMSEA**	**CFI**	**TLI**	**SRMR**
Model 2: Theory-Driven AMPLIFY Agency	1490.750, p=0.000	736	0.04	0.917	0.912	0.062
Model 3: Data-Driven Model of Agency	902.867, p=0.000 n=651	346	0.050	0.909	0.901	0.061

Although neither model passes the exact fit test with a nonsignificant chi square value test, the chi-square statistic is sensitive to sample size and data structure, and accordingly we agree with those who argue that the strength of any SEM should rely foremost on the strength of its conceptual and theoretical underpinnings (Vandenberg, 2006). Therefore, we adopt model 2 as the preferred model, based on both the strength of the conceptual framework derived from literature-informed theory and practitioner input, and the comparative approximate fit statistics which indicate that

model 2 also provides an adequate description of the sample data, and a better fit than model 3. Figure 5 depicts our final SEM of the theory-driven model illustrating four first order factors: self-governance skills (sg), leadership skills (ls), self-beliefs (sb), and environmental beliefs (eb), which combine to form a general factor agency (ag) at the second order level.

Figure 5 *SEM Model 2-AMPLIFY Theory-Driven Agency Framework*



Invariance Analysis

Table 5 depicts the results of a multi-group comparison of the first order factors of the theory-driven SEM. We were not able to establish measurement invariance of the second order factor (agency) as the multigroup second order latent models did not converge. Evaluating the first order factor models, we establish configural and metric invariance as there were no significant differences between the configural and metric models (chi square difference test P value=0.2339). Results also depict significant chi square differences between the scalar and metric models with small CFI differences of less than 0.010 between all models (Chen, 2007). We therefore conclude partial invariance of the measure and confirm a common factorial structure up to the first order latent factors across pre-post groups. This suggests that the agency scale demonstrates structural and measurement invariance up to the first order latent factors, but more evidence supporting invariance of the second order factor structure across groups is required.

Table 5 *Multi-Group First-Order Factor Models*

Model	**# of Parameters**	**Degrees of Freedom**	**Chi-square**	**P-value**	**CFI**
Configural	384	2229.259	1468	0.00	0.910
Metric	348	2218.765	1504	0.00	0.915
Scalar	246	2307.699	1606	0.00	0.917
Models Compared	**Degrees of Freedom**	**Chi-square**	**P-Value**		
Metric against Configural	36	41.787	0.2339		
Scalar against Configural	138	171.007	0.0296		
Scalar against Metric	102	145.329	0.0031		

7. Discussion

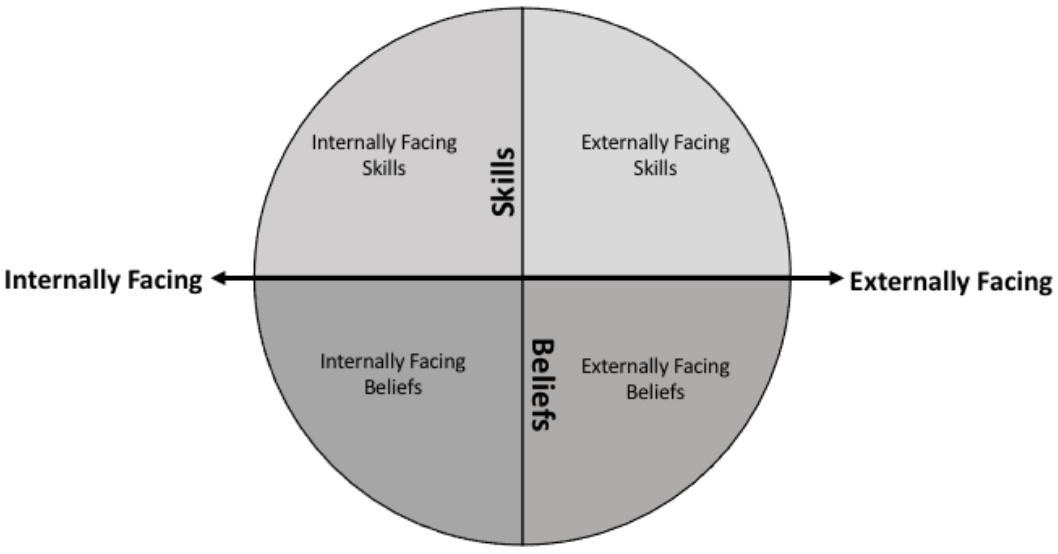
The primary purpose of this study was to develop and confirm a robust practitioner-informed conceptual framework and associated measure of adolescent girls' agency in East Africa. The second aim of the study was to validate our measure of agency for intended use as an evaluative tool assessing the impact of informal life skills education programs across pre/post intervention groups. Results of CFA model comparison suggest that the practitioner-proposed and theory-driven model of agency presents the best fit of our data, indicating that agency for the study population can be viewed as four distinct factors: self-beliefs, environmental beliefs, self-governance skills, and leadership skills. Furthermore, we found no evidence of pre/post group differences in factor structure of our measure of agency across intervention groups up to the first order latent factor scores, indicating that our instrument similarly measures the underlying traits of self-governance, leadership skills, environmental beliefs, and self-beliefs, from baseline to endline, and that pre-post group comparisons of these traits are meaningful. However, more work is needed to establish group comparison of associated agency scores.

The evidence presented herein is a good starting point for understanding agency as a 4-dimensional construct and provides an approach to measuring agency that is missing from the current field—one that "systematically adapts measures to local context" (Donald et al., 2017, p. 52). The involvement of implementers in developing a conceptual framework of agency that represents their context as well as their engagement in survey development is precisely what Galloway et al. (2017) argue is lacking from other existing measures of soft skills. We present this process as an important contribution for how to do the work of measurement in a way that is responsive to context and inclusive of practitioner input.

Four Domains of Agency: An Adaptable Framework for Measurement

Most importantly, we believe that the presented collaborative work has provided a practical tool for understanding agency as a context-dependent construct, and for understanding how it might be adapted for broader populations and different settings. Extrapolating from the conceptual work of the researchers' collaboration with AMPLIFY Figure 6 depicts a generalized framework of girls' agency as four domains of externally- and internally-facing skills and beliefs. Broadly, the framework shows girls' agency as four domains: internally-facing beliefs about her own ability; externally-facing beliefs about the malleability of her environment (including gender norms) and the extent to which these can be negotiated; the internally-facing skills required to organize herself and coordinate her actions for social change; and, finally, the externally-facing skills needed to effect change outside of herself through her individual relationships and personal "power to" (Kabeer, 1999). We believe that this is a useful conceptual tool, consistent with the literature, that provides a basis for *how* practitioners might adapt measures of agency across settings while maintaining conceptual comparability.

Figure 6 *Four Dimensions of Agency*



The AMPLIFY agency measure provides one example of such adaptation. Specifically, the environmental beliefs category of the AMPLIFY agency measure captures a girl's understanding of the rigidity of her environment in East Africa, and importantly the gender norms which structure her opportunities—what Kabeer might have called "oppositional force." This dimension of agency is necessary for understanding agentic capacity as dependent on the intersection of girls' experiences with gender and poverty. Relatedly, AMPLIFY practitioners specified externally-facing skills as leadership skills related to girls' ability to influence the opinions and actions of those around her (evidenced by items in this factor related to public speaking, conflict resolution and interpersonal communication). In other settings with different constraints, it is easy to envision that agency might be dependent on a different set of externally-facing skills—those related to negotiation for example, or local economic opportunities.

In summary, our tool development process suggests an alternative approach to thinking about regional versus universal measures of soft skills. If we acknowledge that soft skills like agency are socio-culturally and socio-economically mediated constructs, then any associated measurement tool will have regional, or even country, limits to its validity. Rather than looking for uniform metrics for soft skills, our approach suggests that quality education experts should rely on regionally specific measures that are conceptually comparable and should test how widely these measures may be valid. Our study offers one example of how to achieve this approach to regional adaptation of soft skill measures.

Epistemological Position—a Practical Undertaking

We acknowledge the inherent contradiction in developing a standardized measure for what we have also defined as a socio-economic and socio-culturally mediated construct. Our position is a practical one, stemming from the very real need and stated desire of the practitioners in the study for a comparable measure of agency that can be used across the region and for achieving consensus

from place to place, recognizing that some important constraints on girls' lives are common across the region, or at least common for girls from similar socio-economic backgrounds. Thus we understand the AMPLIFY agency survey as an instrument whose interpretation is likely to have limits to validity, but which has value for evaluating girls' outcomes regionally and thus can offer a practical (albeit imperfect) solution to an identified problem of practice. We urge further research into the validity of our instrument's interpretation and use within each country to better understand how the dimensions of agency might shift from place to place.

8. Limitations

Although we included the voices and expertise of practitioners in our conceptualization of agency, we admittedly failed to include the voices of adolescent girls. While we are confident that the construction of our conceptual framework accurately represents the views of practitioners and their program teams, a lot could be learned from including girls themselves in both the conceptualization and the measurement of agency and is a priority for further research. Related to this, we were not able to include cognitive interviewing of participants on either the agency construct or the translated tools. Although our instruments were back translated, the translations warrant an independent investigation of interpretation and understanding.

Another important limitation of the AMPLIFY's agency survey, is that it consists of only self-reported measures, which are known to suffer from bias. However, we pursued a self-report structure for our instrument because of the difficulty in developing, testing, and validating observational measures of complex skills, and implementing them reliably in low-resource settings. We recommend that more research be dedicated to developing observational measures that can be used to triangulate data obtained from self-reported instruments in settings like East Africa, and we are keenly following the few such efforts currently underway (Mugo, 2020).

Notably, our data has a number of limitations which restrict the inferences that can be made from our analysis. Primarily, our data is not representative of adolescent girls in East Africa, but rather is only representative of the population served by AMPLIFY's member organizations, limiting the generalizability of our results to the AMPLIFY constituency or similar populations. Furthermore, our sample had a large amount of missing data, thus limiting our ability to conclusively draw inferences from exploratory factor analysis for our data-driven model. Finally, data was skewed by country with the largest sample coming from Kenya preventing us from understanding the potential variation of the measure across country. Our measure also requires a thorough investigation of reliability which was beyond the scope of this study and our data.

Finally, our survey development did not explicitly explore locally grounded understandings of agency by country or community. During the content mapping of the survey, there was no consensus among AMPLIFY members that all skills on the survey were *necessary* for girls' agency; rather, researchers selected those skills that the majority of members identified as *important.* In this way, the AMPLIFY agency survey is an exercise in maximizing commonality, but it is not perfectly representative of any one organization or community's perspective. On the other hand, we hope that the conceptual clarity of the four dimensions of agency will remain a useful tool for conceptualizing agency across place compared to more narrowly defined measures.

9. Conclusion

Our research offers an important contribution to the measurement of agency for adolescent girls. The AMPLIFY agency survey improves on existing measures by assessing agency as a multi-dimensional (as opposed to one-dimensional) construct, and by including a robust practitioner-driven conceptualization of agency for the East Africa region. It further develops this measure specifically for adolescent girls as a unique population with unique constraints on their agencies, where previous efforts have largely focused on adult women. We present a process of survey

development which relies on practitioner collaboration and situates the skills of interest as context-dependent—providing a conceptual framework for how to achieve regional or country-based adaptations of measures of agency. Of the most important priorities for further study we highlight the need for more rigorous investigation into the psychometric properties of the tool proposed, and the regional limits of the survey's interpretation and use as a pre/post measure of adolescent girls' agency.

References

Ahearn, Laura M. (2010). Agency and Language. *Society and Language Use,* 7, 28–48.

AMPLIFY. (2019). Measuring Girls' Agency in East Africa—Co-Creating Contextually Specific Tools for Evaluation: Lessons from the AMPLIFY Girls Collective. AMPLIFY Girls, Nairobi Kenya.

Banati, P., Rumble, L., Jones, N., & Hendriks, S. (2021). Agency and Empowerment for Adolescent Girls: An Intentional Approach to Policy and Programming. *Journal of Youth Development,* 16 (2–3), 239–254.

Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., & Sulaiman, M. (2014). "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa. London School of Economics: STICERD Publications.

Bandura, A. (2006). Towards a Psychology of Human Agency. *Perspectives on Psychological Science,* 1(2), 164–80.

Bandura, A. (2009). Cultivate Self-Efficacy for Personal and Organizational Effectiveness--Indispensable Knowledge for Evidence-Based Management. In E. A. Locke (Ed.), *Handbook of Principles of Organizational Behavior*. West Sussex, UK: John Wiley and Sons, Ltd.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606.

Brush, K. E., Jones, S. M., Bailey, R., Nelson, B., Raisch, N., & Meland, E. (2022). Social and Emotional Learning: From Conceptualization to Practical Application in a Global Context. In *Life Skills Education for Youth* (pp. 43–71). Springer.

Burnett, N., & Jayaram, S. (2012). Skills for Employability in Africa and Asia. ISESE Skills Synthesis Paper. Washington, DC: R4D.

Care, E., & Anderson, K. (2016). *How Education Systems Approach Breadth of Skills* (Skills for a Changing World). The Brookings Institution, Center for Universal Education. https://www.brookings.edu/wp-content/uploads/2016/05/brookings_how-education-systems-approach-breadth-of-skills_v2-1.pdf.

Catalano, R. F., Skinner, M. L., Alvarado, G., Kapungu, C., Reavley, N., Patton, G., Jessee, C., Plaut, D., Moss, C., Bennett, K., Sawyer, S. M., Sebaney, M., Sexton, M., Olenik, C., & Petroni, S. (2019). Positive youth development programs in low-and middle-income countries: A conceptual framework and systematic review of efficacy. *Journal of Adolescent Health,* 65(1), 15–31.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal* 14(3), 464–504.

C-Change. (2011). *Gender Equitable-Men's Scale,* in Compendium of Gender Scales. http://gender.careinternationalwikis.org/_media/c-change_gender_scales_compendium.pdf

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation,* 10(1), 7.

DeVellis, R. F. (2016). *Scale Development: Theory and Applications*. SAGE Publications, Inc.

Dickson, K., and Bangpan, M. (2012). *Providing Access to Economic Assets for Girls and Young Women in Low-and-Lower Middle-Income Countries: A Systematic Review of the Evidence*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Donald, A., Koolwal, G., Annan, J., Falb, K., & Goldstein, M. (2020). Measuring women's agency. *Feminist Economics*, 26(3), 200–226.

Donald, A., Koolwal, G., Annan, J., Falb, K., & Goldstein, M. (2017). Measuring Women's Agency. Policy Research Working Paper 8148. World Bank Group Africa Region.

Duckworth, A.L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and Passion for Long-Term Goals. *Journal of Personality and Social Psychology,* 92(6), 1087–1101.

Ellis, C.M., and Chaffin, J. (2015). Evaluations of Outcomes for Children and Youth from NGO-Supported Microeconomic Interventions: A Research Synthesis. *Enterprise Development and Microfinance,* 26(2), 104–21.

Erulkar, A., Ferede, A., Girma, W., & Ambelu, W. (2013). Evaluation of 'Biruh Tesfa' (Bright Future) Program for Vulnerable Girls in Ethiopia. *Vulnerable Children and Youth Studies,* 8(2), 182–92.

Galloway, T., L. Lippman, H. Burke, O. Diener, and S. Gates. (2017). Measuring Soft Skills and Life Skills in International Youth Development Programs--A Review and Inventory of Tools. USAID's YouthPower Implementation IDIQ-Task Order 1. YouthPower Action. www.fhi360.org/sites/default/files/media/documents/resource-yp-measuring-soft-skills.pdf.

Heckman, J., Stixrud, J., and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. Working Paper 12006. NBER Working Paper Series. National Bureau of Economic Research.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal,* 6(1), 1–55.

Ibarrarán, P., Ripani, L., Taboada, B., Villa, J. M., & Garcia, B. (2014). Life Skills, Employability, and Training for Disadvantaged Youth: Evidence from a Randomized Evaluation Design. *IZA Journal of Labor and Development*, 3(1), 10.

Jackson, K. (2012). Non-Cognitive Ability, Test Scores, and Teacher Quality: Evidence from 9th Grade Teachers in North Carolina. Working Paper 18624. NBER Working Paper Series. National Bureau of Economic Research.

Jayachandran, S., Biradavolu, M., & Cooper, J. (2021). *Using machine learning and qualitative interviews to design a five-question women's agency index*. National Bureau of Economic Research. https://www.nber.org/papers/w28626.

Kabeer, N. (1999). Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment. *Development and Change,* 30(3), 435–64.

Kane, M. T. (2013). Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement*, 50(1), 1–73.

Kane, M., & Trochim, W.M.K. (2007). *Concept Mapping for Planning and Evaluation* (Vol. 50). Sage Publications, Inc.

Kautz, T., Heckman, J. J., Diris, R., ter Weel, B., & Borghans, L. (2014). Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success. Working Paper 20749. National Bureau of Economic Research.

Learning Metrics Task Force. (2013). Toward Universal Learning: A Global Framework for Measuring Learning. Report 2 of the Learning Metrics Task Force. UNESCO and The Brookings Institution.

Leventhal, K.S., Gillham, J., DeMaria, L., Andrew, G., Peabody, J., & Leventhal, S. (2015). Building Psychosocial Assets and Wellbeing among Adolescent Girls: A Randomized Controlled Trial. *Journal of Adolescence,* 45(December), 284–95.

Lippman, L. H., Ryberg, R., Carney, R., & Moore, K.A. (2015). Workforce Connections: Key "soft skills" that foster youth workforce success: Toward a consensus across fields. Washington, DC: Child Trends.

Little, T., Snyder, R., & Wehmeyer, M. (2006). The Agentic Self: On the Nature and Origins of Personal Agency Across the Lifespan. In D. K. Mroczek & T. D. Little (Eds.), *The Handbook of Personal Development*. Lawrence Erlbaum Associates.

Lloyd, C. B., & Hewett, P. (2009). Educational Inequalities in the Midst of Persistent Poverty: Diversity Across Africa in Educational Outcomes. *Journal of International Development,* 21, 1137–1151.

Lopez, A., Yoder, J. R., Brisson, D., Lechuga-Pena, S., & Jenson, J. M. (2015). Development and Validation of a Positive Youth Development Measure: The Bridge-Positive Youth Development. *Research on Social Work Practice,* 25(6), 726–36.

Mincemoyer, C. C., & Perkins, D.F. (2005). Measuring the Impact of Youth Development Programs: A National On-Line Youth Life Skills Evaluation System. *The Forum for Family and Consumer Issues,* 10(2), 1–9.

Mugo, J. (2020). Welcome ALiVE—RELI's Firstborn Initiative (Assessment of Life Skills and Values in East Africa). Regional Education Learning Initiative. https://reliafrica.org/welcome-alive-relis-firstborn-initiative/

Mulaik, S. (2007). There Is a Place for Approximate Fit in Structural Equation Modelling. *Personality and Individual Differences,* 42(5), 883–91.

Murphy-Graham, E., & Lloyd, C. (2016). Empowering Adolescent Girls in Developing Countries: The Potential Role of Education. *Policy Futures in Education*, 14(5), 556–77.

Murphy-Graham, E., & Leal, G. (2015). Child Marriage, Agency, and Schooling in Rural Honduras. *Comparative Education Review,* 59(1), 24–49.

Muthén, L. K., & Muthén, B.O. (2017). *Mplus User's Guide*. 8th ed.

Nasheeda, A., Abdullah, H.B., Krauss, S.E., & Ahmed N.B. (2019). A narrative systematic review of life skills education: Effectiveness, research gaps and priorities. *International Journal of Adolescence and Youth,* 2(3), 362–379.

Neal, S. E., Chandra-Mouli, V., & Chou, D. (2015). Adolescent first births in East Africa: Disaggregating characteristics, trends and determinants. *Reproductive Health,* 12(1), 1–13.

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global Self-Esteem and Specific Self-Esteem: Different Concepts, Different Outcomes. *American Sociological Review,* 60(1), 141–56.

Scheier, M.F., & Carver, C. S. (1985). The Self-Consciousness Scale: A Revised Version for Use with General Populations. *Journal of Applied Social Psychology,* 15(8), 687–99.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research Online,* 8(2), 23–74.

Scholz, U., Doña, B. G., Sud, S., & Schwarzer, R. (2002). Is General Self-Efficacy a Universal Construct? Psychometric Findings from 25 Countries. *European Journal of Psychological Assessment,* 18(3), 242.

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy Scale. In J. Weinman, S. Wright, and M. Johnston (Eds.), *Measures in Health Psychology: A User's Portfolio*. NFER - Nelson.

Sidle. (2019). Action on Agency: A Theoretical Framework for Designing and Operationalizing Agency in Girls' Life Skills Programs. *Gendered Perspectives in Development Working Paper Series*, #313 Michigan State University.

Sireci, S. G. (2016). On the validity of useless tests. *Assessment in Education: Principles, Policy & Practice*, 23(2), 226–235.

Torre, D. A. (1986). *Empowerment: Structured Conceptualization and Instrument Development*. Cornell University.

Temin, M., & Heck. C. J. (2020). Close to Home: Evidence on the Impact of Community-Based Girl Groups. *Global Health: Science and Practice,* 8(2), 300.

UNESCO. (2021). *Global education monitoring report, 2021/2: Non-state actors in education: Who chooses? Who loses?* UNESCO. https://www.education-inequalities.org/indicators

Vandenberg, R.J. (2006). Statistical and Methodological Myths and Urban Legends: Where, Pray Tell, Did They Get This Idea? *Organizational Research Methods,* 9(2), 194–201.

Wado, Y. D., Sully, E.A., & Mumah J.N. (2019). Pregnancy and early motherhood among adolescents in five East African countries: A multi-level analysis of risk and protective factors. *BMC Pregnancy and Childbirth,* 19(1), 1–11.

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods,* 51(1), 409–428.

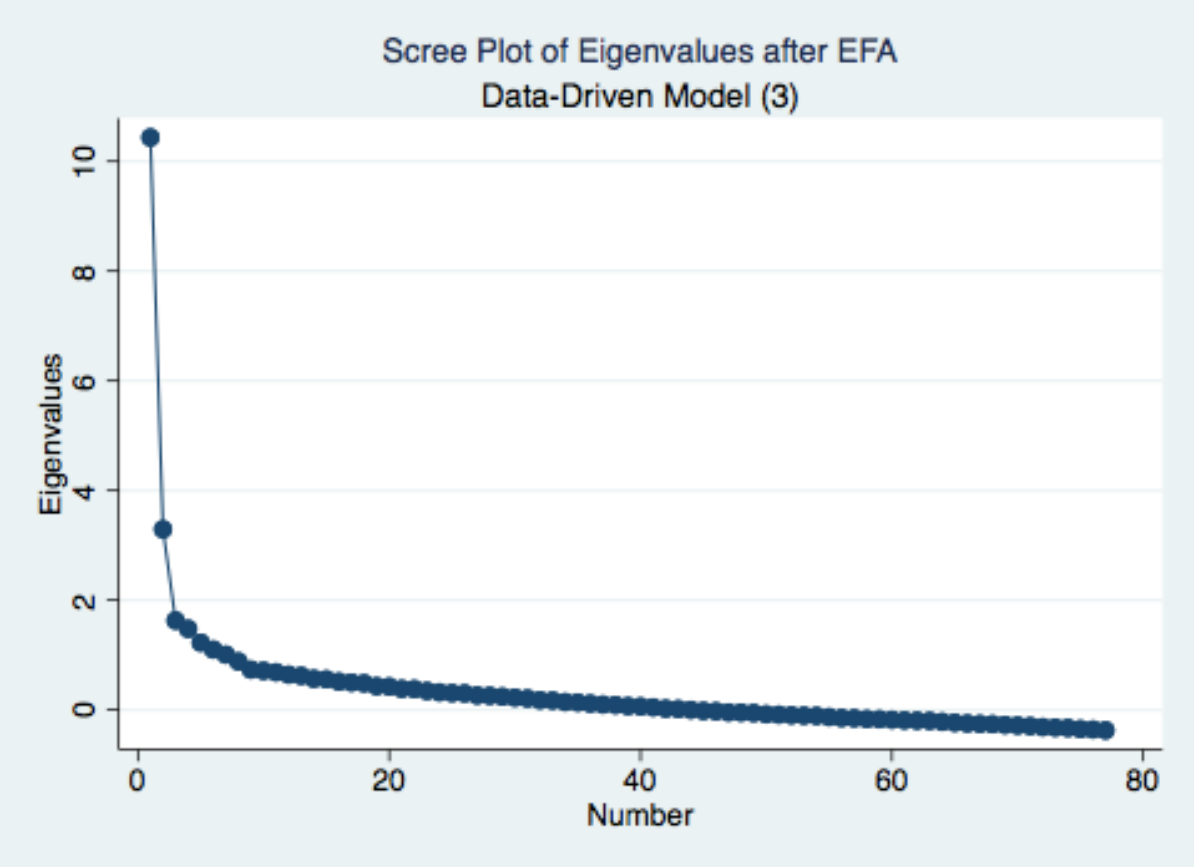
Zins, Joseph E., Weissberg, R. P., Wang, M. C., & Walberg, H. J. (2004). *Building Academic Success on Social and Emotional Learning: What Does the Research Say?* Teachers College Press.

Appendix:

Table 3: Factor Correlation Matrix of the Four Data-Driven Derived Factors

	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**
Factor 1	1			
Factor 2	0.4112	1		
Factor 3	0.2317	0.2771	1	
Factor 4	0.1034	0.07217	0.1092	1

Figure 4: Scree plot of Eigenvalues after EFA: Data-Driven Model (3)



PAPER #3: TOWARDS A THEORY OF BEST PRACTICE FOR GIRLS' AGENCY IN EAST AFRICA

Aubryn Allyn Sidle

Abstract: Despite numerous experimentally-tested programs in girls' education, there is no consensus on best practice for improving girls' outcomes. In this paper I argue that instead of seeking to scale evidence-based interventions, the sector should identify scalable *theories* of practice. Such theories necessitate a shift away from attempting to replicate successful programs irrespective of context, toward designing programs *based* on context in order to replicate *outcomes*. I use Group Concept Mapping as a participatory tool for group conceptualization and theory-building, in collaboration with the AMPLIFY Collective of community-driven organizations (CDOs), to construct a theory of practice for improving adolescent girls' agency—one that is empirically driven and collaboratively built by practitioner and CDO expertise.

Introduction: Girls' Education—A Sector in Need of a Theory

The question of identifying and scaling best practices in girls' education has become more pressing in the wake of the Covid-19 pandemic and the related crisis in girls' secondary school dropout rates (Oulo et al., 2021, Presidential Policy and Strategy Unit (Kenya) Population Council, 2021). A handful of recent publications have attempted to map evidence-based interventions and curricula for "what works" to support girls, focusing on programs that show promise for "scale" and "replicability" (Psaki et al., 2021; Sridharan et al., 2019; Sperling & Winthrop, 2016). Scale, as it is currently imagined, involves identifying experimentally-tested interventions and then replicating them across broad geographies. In the culturally and institutionally diverse regions of Eastern Africa, this seems like an unlikely strategy.

A more reasonable approach to scaling, proposed in this paper, is to consider the practices and processes most likely to replicate an outcome in a particular environment, as opposed to replicating a specific intervention *irrespective* of environment. This approach necessitates an intellectual shift from considering programs as the unit to be scaled, toward thinking about practices and outcomes. The girls' education sector does not need more evidence-

based programs. Many of these programs exist, but they have failed to proliferate. Instead, the sector needs practitioner-informed *theories* of practice that re-frame the idea of 'scalability' to include robust considerations of context and clear hypotheses about the causal processes that produce or hinder a desired result. Such context-inclusive theories would provide a better guide to practitioners for replicating success across diverse settings.

The purpose of this paper is to construct a working theory of practice as a heuristic tool for how to design programs that replicate girls' outcomes across diverse communities in East Africa. I argue that most programs in girls' education fail to account for environment, and thus replicating them fails to replicate outcomes. Furthermore, the rationale for scaling "evidence-based" practices is based on experimental evaluations which too often conflate internal with external validity. I propose a different approach to generating a theory of best-practice—one which attends to the individual *as well as* to the individual's environment, is driven by community-based practitioners, and which approaches the question of generalizability through heterogeneity rather than randomization.

I first make an argument for why such "context-inclusive" theories of practice are needed for improving youth outcomes. Secondly, I introduce my research collaboration with AMPLIFY Girls, identify "agency" as the outcome of interest for this study, and discuss how and why community-based organizations offer an important perspective on the question of "best" practice. Next, I attempt to answer the question "*what are best practices for producing girls' agency?"* through a review of the practice literature triangulated with analysis of empirical evidence from 18 community-driven organizations and their private funders. Utilizing Group Concept Mapping (Kane & Trochim, 2007) as a participatory tool of theory building (Trochim, 1985), this

evidence is analyzed with the purpose of generating a *theory* of best practice for improving adolescent girls' agency.

The Premise of Scalability and the Omission of Context

Randomized Control Trials (RCTs) have been considered the "gold standard" of evidence in education since the 1990s (Maxwell, 2004; Biesta, 2007; Scriven, 2008). Most experimental studies evaluate programs as uniform entities, seeking to estimate with a high degree of internal validity the effect size of a program on clearly defined and measurable outcomes. If the study population is appropriately sampled and randomly assigned to either a treatment or a control group, then the assumption of the RCT methodology is that the study findings are generalizable to similar populations.

However, assessing what constitutes "similar" in population composition is a complicated question. There is mounting evidence to suggest that researchers should not generalize RCT findings, especially across country or regional borders (Hanushek, 2021), or target populations in settings different from the experiment (Stuart et al., 2015). In his 2021 review of education impact evaluations, economist Eric Hanushek found that the diversity of institutional settings across countries, combined with limited evidence replicating results in other contexts, undermined the generalizability of RCTs, making "attempts to generalize results across countries appear very risky" (Hanushek, 2021, p. 9). Hanushek's conclusions echo earlier demographic research suggesting that contextual factors such as the "etiology" of school dropout, the norms around a young women's life course, and the stage of demographic transition in a particular country, all affect the impact of girls' education investments (Eloundou-Enyegue & Stokes, 2004).

In a recent experimental evaluation of BRAC's Empowerment and Livelihood for Adolescents (ELA) program, substantial effects were found on youth outcomes in Uganda but not Tanzania. The authors largely attribute this to differences in implementation fidelity, although they reference other "contextual factors" (Buehran et al., 2017, p. 12). The fact that such conclusions about contextual influence appear to be novel underscores a fundamental assumption of the sector: that programs developed and successfully tested with low-income girls in one country (such as Ethiopia) are generalizable to another (like Tanzania). Such assumptions effectively equate the material aspects of girls' lives in both countries despite vastly different cultural, ethnic, linguistic, economic, and social settings. Yet systematic reviews of girls' education program evaluations continue to cite experimental results as definitive evidence for scaling up programs to multiple countries as being best practice (Psaki et al., 2021; Catalano et al., 2019; Sridharan et al., 2019).

I agree with Murphy-Graham and Leal, who argue that the socio-cultural and economic environments of each community produce a diversity of structural constraints that tangibly affect the outcomes of educational programs (2015). The way girls experience outcomes like agency, for example, is heavily influenced by their experiences of poverty and the various opportunity structures in their communities (Klocker, 2007). Calls for evidence-based scale often conflate internal validity with external and ignore the practical fact that current girls' education programming is causally complex, structurally multi-component, and context-specific in both content and design.

Recent analysis of the largest global database of girls' education programs shows that 63% of programs include five or more components, compared to just 11% of programs with just one component (Psaki et al., 2021, p. 47). The most common interventions found in this database

were informal life skills education curricula delivered through after-school girls' clubs. These initiatives have complex causal mechanisms with primary interventions meant to directly impact a variety of soft skills (e.g. "self-confidence" or "empowerment"), that in turn are intended to improve educational achievement, long-run health outcomes, and economic well-being (Benson et al., 2011; Catalano et al., 2004). Such causal complexity is notoriously hard to evaluate, let alone replicate. I argue that the multi-component nature of many programs is not simply a failure of the sector to focus strategically on "what works"; rather it reflects the reality that healthy youth development depends on multiple facets of a girl's environment.

Towards a Context-Inclusive or 'Ecosystem' Approach to a Theory of Practice Theory

Context-inclusive theories of practice offer a more practical approach to replicability and scale because positive outcomes are understood in relationship to girls' environments. Human development scholars have established that youth development unfolds as a *process* in which young people interact with their environments (Bronfenbrenner & Morris, 1994) and are shaped by the social, ecological, and physical structures around them (Lerner et al., 2005). On this view, the theory of practice underlying any program approach must account for the characteristics of a girl's environment (or ecosystem) that might foster or constrain her success.

There are excellent examples of such "ecosystem" theories in youth development. For instance, the developmental assets framework for youth "posits a theoretically-based and research-grounded set of opportunities, experiences and supports that are related to promoting school success, reducing risky behaviors and increasing socially-valued outcomes including prosocial behavior, leadership and resilience" (Benson et al., 2011, p. 197). The developmental assets framework proposes that healthy youth development is dependent on two types of assets:

"internal assets" (including positive skills, competencies, knowledge, and behaviors) and "external assets" (e.g. community, and environmental influences) (Damon & Gregory, 2003).

Benson further expands on the notion of external assets to say that they are comprised of four types of infrastructures within a community: economic, service, program, and human development (2003, p. 21–22). Thus, the Developmental Assets Theory implicitly argues that successful youth program design should account for the differing level of external assets that exist in a community, and directly address gaps in infrastructure through interventions. Such a context-inclusive theory offers sound guidance to practitioners, and a practical heuristic device for strategically constructing programs based on specific aspects of context that might be present or absent in a given setting.

The girls' education sector requires similar context-inclusive theories—ones that envisage an ecosystem of practices that will either mitigate constraining aspects of context, and/or enhance enabling ones in order to replicate *outcomes*. Put another way, a context-inclusive theory will not only focus attention on the interventions aimed at producing changes in the *individual,* but it will also include a vision for the program practices that might produce enabling environments (Archibald, 2015). Here I use the term "practices" to indicate the broad types of community and youth engagement, the philosophies about that engagement, the pedagogical strategies, and the content that inform the development of specific interventions.

Background & Theoretical Approach

AMPLIFY Girls

I address this need for a context-inclusive approach to a theory of practice in collaboration with the AMPLIFY Girls Collective of community-driven organizations (CDOs) in East Africa. Founded in 2017, AMPLIFY is a self-governed organization that up until 2020 largely made organizational decisions through consensus of the collective membership. Its membership is

made up of 18 community leaders/practitioners who are the founders and/or Executive Directors of community-driven organizations serving adolescent girls in Rwanda, Kenya, Tanzania, and Uganda. The collective is governed by a board of directors elected from amongst the membership and representing each country in the collective.

AMPLIFY members use the term "community-driven organization" (CDOs) rather than "community-based organization" (CBOs) to acknowledge the fact that in order to succeed in the globalized economy many CBOs have had to establish relationships with sister organizations located in the global North, who serve as channels for directing international financial resources to local programs. Thus "CDO" in the context of AMPLIFY includes organizations who both have and do not have such partnerships overseas, but where power and decision making about programs and practice are made by local leaders who are located in, and who are also from, the community of practice.

In this study I center the expertise and priorities identified by AMPLIFY CDOs as the primary evidence for theorizing best practice. In doing so, I am admittedly assuming that these practitioners effectively represent the diversity of voices of the girls they serve. Obviously, there are limitations to this assumption. However, it is a calculated risk made based on an understanding that CBOs have proved effective at delivering aid broadly (Casey, 2018), that community-based youth organizations are particularly effective at serving "minoritized" and "marginalized" youth (Wells et al., 2021; Baldridge et al., 2017), and that community-based practitioners are usually better qualified than international technical experts to design programs that meet their communities' needs.

It was in 2018 that I began a three-year research collaboration with AMPLIFY to support community leaders' request to develop a shared set of metrics for documenting their impact on

adolescent girls' outcomes. Utilizing Group Concept Mapping, we co-created a shared metrics framework that identified girls' agency as the primary outcome of interest for the collective. Throughout 2019 and 2020, I worked extensively with AMPLIFY community leaders to build each organization's capacity to track agency outcomes, and to improve their understanding of research processes and methods. I also provided training in Group Concept Mapping (GCM) as a research tool. During this period, we built a regional measure of agency to evaluate collective impact. I became very familiar with AMPLIFY organizations' programs, and I made site visits to six AMPLIFY member communities where I spent time with key stakeholders, girls, and recent alumnae. AMPLIFY members became well-versed in the GCM methodology and comfortable collaborating on research studies using this tool. A full write up on these shared metrics and the associated agency instrument can be found in Sidle et al. (2019) and Sidle and Oulo (*forthcoming)*.

Feminist Evaluation Research

My approach to research is driven by the epistemological orientation of feminist evaluation which centers inquiry about women and girls, and views knowledge as "situated" and "relational" (Brisolara, 2014, p. 16). Such a view necessitates reflexivity from the researcher, the privileging of local voices, and an orientation to research that understands knowledge-gathering as a political act (ibid). Accordingly, I situate myself in this study as both researcher and participant, as a former youth development and girls' education practitioner and a current scholar who draws on my own relationships with participants, my knowledge of their programs, and my expertise—both practical and academic—to inform my analysis.

Our research collaboration is political in that it privileges the expertise of leaders from the global South over "technical expertise" in an attempt to reframe who may serve as legitimate sources of knowledge (Brisolara et al., 2014). The resulting collaboration is intended to make a

contribution to policy that suggests an alternative framing of the value of both community-driven and Southern organizations and their approaches to global problems of practice. Our research question was born out of AMPLIFY members' desire to conceptualize a set of program practices and interventions that represent a community-driven vision for supporting girls and, in so-doing, offer an alternative, community-driven framework for scale. Together, we developed and undertook to answer the question: *What are best practices for improving girls' agency?*

Literature Review: The Practices that Improve Agency

A first step toward answering our research question necessitates a review of the practice literature for clues on emergent consensus around practices or programs from the evaluation and experimental research that improve girls' agency. There is no sector-wide consensus on the definition and measurement of agency, although it is usually operationalized in program evaluations as constructs related to girls' "empowerment" (e.g. self-efficacy, or resilience) or autonomous choice (Donald et al., 2020). In my analysis, I use a definition of agency developed by AMPLIFY members, who conceptualize the notion as a combination of internally and externally facing beliefs (about self and environment) and skills (related to leadership and self-governance) which support a girls' change-making capacity (Sidle & Oulo, *forthcoming)*.

On the basis of this definition, my review of the literature considered programs that have been shown to influence a wide range of skills and beliefs related to the ability of a girl to affect change in her life or environment. In some cases the outcome indicator of these programs was behavioral or attitudinal, such as reduction in child marriage or an improvement in gender positive beliefs. In most cases, however, these outcomes were soft skills like self-efficacy, resilience, decision-making, or other related capacities that could be grouped into categories of "self-beliefs" or "self-governance skills." In all cases, outcomes reported in each program were

not fully representative of the multi-dimensional construct of agency, and as such I consider the literature useful in terms of offering clues towards common practices, but not conclusive.

My literature review included over 40 books, articles, and/or reports of evaluations on programs in girls' education, identified using key word search protocols in google scholar and the resource data bases of large policy-institutions such as the World Bank. Of these 40, there were 23 studies that showed evidence of positive effects of the program on specific indicators related to adolescent agency (as defined above). Eighteen of these studies were of girl-specific programming, while five studies focused on youth broadly. Seven studies were systematic reviews of experimental evidence (Temin & Heck, 2020; Catalano et al., 2019) or non-experimental program evaluations (Alavarado et al., 2019; Unterhalter et al., 2014; Catalano et al., 2004) and/or summaries of the findings from multiple interventions (Banati et al., 2021; Sridharan et al., 2019). Altogether, this literature represents insights from more than 400 programs.

I read the relevant literature for evidence of practices (as opposed to specific programs) that have a demonstrated impact on agency-related skills or beliefs. Practices include engagement strategies, delivery mechanisms, pedagogical approaches, and types of content that are associated with successful agency outcomes. In most cases the description of interventions evaluated in these studies was quite short, making the identification of practices a rather straight-forward task. Also, programs with multiple components were mostly evaluated as a whole, and so I was unable to parse which specific practices in a multi-component program were determined to have led to the agency outcome. As previously noted, this is a common challenge with the program evaluation literature as whole.

Ultimately, I found four common practices (found in 13–14 studies), and a fifth practice found in nine out of the 23 studies that I highlight below.

1. **Peer communities of support**: programs that create peer communities of support, putting youth in relationship and dialogue with one another about their lives and challenges.
2. **Soft-skills curriculum:** soft skills curriculum that is structured (both in delivery and content) and is *facilitated* by a trained mentor (peer or adult).
3. **Student-centered pedagogy:** especially project-based learning and/or hands-on opportunities for youth engagement in their communities.
4. **Positive female role model**: programs offering girls the opportunity to build relationships and engage with a positive female role model (adult or peer).
5. **Critical pedagogy**: Programs that utilize critical pedagogy as characterized by the practice of critical dialogue to support girls' self and social-awareness, critical thinking and positive identity formation.

In addition to the five program-practice areas, two other common characteristics of successful programs also emerged from the review. First was attention to multiple aspects of girls' well-being (e.g. psychosocial well-being and economic well-being) as evidenced by engagement with multiple curricular areas (e.g. soft skills and financial literacy) and/or provision of interventions that address girls economic, health and/or psychosocial needs (Banati et al., 2021; Temin & Heck, 2020; Amin et al., 2016; Austrian et al., 2016; Bandiera, 2014). The second characteristic was engagement with multiple stakeholders and/or multiple social spheres (e.g. work in schools and homes, or engagement with girls and teachers or parents) (Temin & Heck, 2020; Alvarado et al., 2017; Murphy-Graham & Leal, 2015; Catalano et al., 2004).

Table 1 (below) displays this synthesis of the literature—showing each practice area along with supportive studies and their brief description. Studies not cited under a specific practice area were excluded because the level of detail of program structure provided in the study was insufficient to determine whether or not it fit the category. Therefore, excluded studies and programs may also share the practice areas outlined here.

This review provides a good foundation for understanding successful program practice for girls' agency as being holistic in nature and comprehensive in approach. We can see a pattern of practices emerging (peer support, soft skills curriculum, role models, alternative pedagogies) that are likely successful for influencing the individual-level change, coupled with an emerging view that holistic engagement with a girls' social setting might be important for her agentic development. The literature does not, however, offer a clear framework for understanding a set of practices and community engagement as necessary for producing agency in East Africa. For a more contextualized view of girls' program practice, I turn to empirical evidence.

Table 1: Literature Review

Practice Area	Empirical studies that support specific practice area for promoting girls' agency.	Year & Geographic Focus	Total # of studies	Description of study
Peer Communities of Support	Banati, Rumble, Jones, & Hendriks (2021)	2021 Global	14	A summary of the evidence from 3 girls programs in Bangladesh (BALIKA), Jordan (Makani) and East Africa (E+E=E)
	Temin & Heck (2020)	2020 Global		Systematic review of 30 experimental and quasi-experimental evaluations of girl groups in LMICs
	Willan, Gibbs, Shai, Ntini, Petersen, & Jewkes (2020)	2020 South Africa		Qualitative evaluation of the Stepping Stones Creating Futures program in South Africa.
	Catalano, Skinner, Alvarado, Kapungu, Reavley et al. (2019)	2019 Global		Systematic review of 35 experimental and quasi-experimental evaluations of positive youth development programs (PYDs) in LMICs.
	Alvarado, Kapungu, Moss, Plaut, Reavley, & Skinner (2017)	2017 Global		A systematic review of 94 program evaluations of Positive Youth Development programs in LMICs.
	Amin, Ahmed, Saha, Hossain, & Haque (2016)	2016 Bangladesh		Results from the RCT of the BALIKA Life Skills project in Bangladesh and its effect on child marriage rates.
	Austrian, Hewett, Soler-Hampejsek, Bozzani, Behrman, & Digitale (2016)	2016 Africa		Midline results on the effects of the Adolescent Girls Empowerment Program RCT in Zambia on girls' empowerment (incl. self-efficacy, gender beliefs & health knowledge).
	Dierker (2016)	2016 US		A case study of you agency in a Youth Space PYD program amongst African American youth in the US.
	Leventhal, K.S., DeMaria, Gillham, Andrew, Peabody & Leventhal S. (2016)	2016 India		Results from a randomized control trial on the effects of Girls' First program (India) on education and psychosocial outcomes.
	Poteat, Calzo & Yoshikawa (2016)	2016 US		Evaluation of Gay Straight Alliance clubs in Massachusetts high schools for promoting youth well-being.
	Shah, DeJaeghere, Josic, & McCleary (2016)	2016 India		An examination of CARE's KGBV program for promoting gender equality and girls' empowerment in Gujarat, India.
	Cervantes-Soon (2017)	2017 Mexico		An ethnographic study of girls' agency in an alternative high school in Jaurez, Mexico.
	Sahni (2017)	2017 India		A qualitative case study of the Prerna Girls' school in India.
	Unterhalter, Arnot, Lloyd, Moletsane, Murphy-Graham, North, Parkes & Saito (2014)	2014 Global		A systematic review of 169 girl-focused programs and their effects on girls' education and gender equality.
			14	See above.

Soft skills curriculum	Banati et al. 2021, Temin & Heck 2020, Willan et al. 2020, Catalano et al. 2019, Alvarado et al. 2017, Amin et al. 2016, Austrian et al. 2016, Dierker 2016, Leventhal et al. 2016, Poteat et al. 2016, Shah et al. 2016, Sahni 2017			
	Sridharan, Ravindranath, & Cosentino (2019)	2019 Africa & India		An evaluation of the 23 programs in the PSIPSE girls' life skills portfolio in sub-Saharan Africa and India
	Catalano, Berglund, Ryan, Lonczak, & Hawkins (2004)	2004 US		A systematic review of 77 Positive Youth Development (PYD) programs in the US and their effect on PYD outcomes.
Positive Female Role Model	Banati et al. 2021, Temin & Heck 2020, Willan et al. 2020, Catalano et al. 2019, Alvarado et al. 2017, Amin et al. 2016, Austrian et al. 2016, Dierker 2016, Leventhal et al. 2016, Poteat et al. 2016, Shah et al. 2016, Sahni 2017		**13**	See above.
	Hull, Jury, Sahni, Kirshner & Middaugh (2015)	2015 India		A qualitative study of a feminist media program and its impact on girls' agency in India.
Student-centered pedagogy	Banati et al. 2021, Willan et al. 2020, Catalano et al. 2019, Alvarado et al. 2017, Dierker 2016, Leventhal et al. 2016, Poteat et al. 2016, Shah et al. 2016, Sahni 2017, Hull et al. 2015, Catalano et al. 2004		**13**	See above.
	DeJaeghere & Lee (2011)	2011 Bangladesh		An investigation of the conditions of marginalization and empowerment in school settings in Bangladesh.
	Kwon (2006)	2006 US		A study of youth activism amongst Bay Area youth of color in the US.
Critical Pedagogy	Willan et al. 2020, Dierker 2016, DeJaeghere & Lee 2011, Kwon 2006, Sahni 2017 and Hull et al. 2015.		**9**	See above.

	Bajaj (2015)	2015 Global		A study of critical peace education that draws on 3 empirical case studies in Mexico, India and the US.
	Morris (2019)	2019 US		A series of qualitative case studies of program practices to promote Black Girls' healing and well-being in the US.
Engagement with Multiple Stake holders	Banati et al. 2021, Temin & Heck 2020, Alvarado et al 2017, Amin et al. 2016, Austrian et al. 2016, Catalano et al. 2004		**7**	See above.
	Murphy-Graham & Leal (2015)	2015 Honduras		Qualitative longitudinal study of girls' agency in Honduras
Addressing Multiple Facets of Well-Being (e.g. psychosocial + Economic)	Banati et al. 2021, Temin & Heck 2020, Amin et al. 2016, Austrian et al. 2016,		**5**	See above.
	Bandiera (2014)	2014 Uganda		Randomized Control Trial of a girls' empowerment program in Uganda that coupled vocational training with SRH education.

Methodology:

Group Concept Mapping as a Tool for Participatory Theory-Building

We use Group Concept Mapping (GCM) for our empirical, theory-building exercise. This participatory tool "enables a group of people to articulate and depict graphically a coherent conceptual framework or model of any topic or issue of interest" (Trochim & McLinden, 2016, p.1; Kane & Trochim, 2007). Utilizing qualitative data generated from brainstormed statements, unstructured sort data, and rate data, GCM produces a visual geography or conceptualization of group consensus on key thematic categories (Kane & Trochim, 2007). Sort data is analyzed using multi-dimensional scaling and hierarchical cluster analysis to produce the visualization of group consensus, and these are then overlayed with the participant rate data.

Each stage of data collection and analysis in GCM can be done in person or remotely, and it can be either facilitated by the researcher or autonomously completed. Analysis produces multiple potential "solutions" to the concept map which are intended to be discussed and revised collaboratively with participants (Kane & Trochim, 2007). For AMPLIFY, this flexibility, ease of engagement, and collaborative structure made concept mapping an accessible tool that could effectively engage multiple stakeholders with a wide diversity of technical and technological skills, and across geographic locations during a time of social distancing, international travel bans, and isolation.

Although mostly used for evaluation, concept mapping is also useful for theory-building (Trochim & Linton, 1986) with demonstrated use for identifying community-driven programmatic interventions (Vaughn et al., 2013). For our purposes, we use GCM as a conceptualization method for building a theory of program practice that articulates the pattern of necessary program practices for producing an outcome (Trochim & Linton, 1985, p. 584). This

application of GCM to theory, "should be based on *both* the consistency of the theory with social science knowledge and with the experiences and conjectures of constituencies that are directly familiar with the phenomenon in question" (Trochim & Linton, 1985, p. 586). To address these requirements, I construct our theory using data derived from key stakeholders of the AMPLIFY collective (and primarily community leaders), and then triangulate their expertise with what we now know from the practice literature about programs that improve girls' agency.

External Validity

Group Concept Mapping is also unique because it facilitates group conceptualization across a wide diversity of participant backgrounds. Such sampling for "heterogeneity" is one way to address the problem of cross-country generalizability. Cook and Campbell urge differentiation between generalizing "to" a well-specified population versus "across" populations in experimental research, warning that studies can easily conflate the former with the latter (1979, p. 71). In East Africa, only an internationally representative sample and related sub-population analysis would achieve the ability to generalize experimental findings across national borders. Such a sample size is practically impossible to achieve, both because resources are scarce and also because programs in girls' education rarely reach such a large population.

Instead, researchers and practitioners need to look to alternative methods for bolstering external validity through proximal similarity modelling (Trochim, Donnelly, & Arora, 2016, p. 84) and purposive sampling strategies (Cook & Campbell, 1979). In our case we use Group Concept Mapping to intentionally represent the lines of diversity found across the region, whether they are institutional (Hanushek, 2021), socio-cultural (Murphy-Graham & Leal, 2015), or contextual (Klocker, 2016). We achieve this through the diversity of the AMPLIFY membership whose communities are located in four countries with distinct institutional settings, and whose populations represent rural, urban, and peri-urban communities that are engaged in

farming, pastoralism, and/or casual day labor, with a wide representation of religious beliefs (Muslim, Christian, and indigenous) and ethnic and linguistic groups.

Participant Description

Participants fall into one of four categories of key stakeholders: AMPLIFY members (founders/executive directors of AMPLIFY partner organizations and/or their lead program staff), AMPLIFY staff (employees of the AMPLIFY Girls organization), AMPLIFY donors (both small and large private foundations based in the US but operating in multiple countries) and the researcher (myself). AMPLIFY organizations determined who would serve as a "participant" from each of their teams, while AMPLIFY's fulltime staff (three participants) and the researcher (at members' request) were also invited to participate. AMPLIFY staff identified a small cohort of AMPLIFY's most important donors and invited them to participate also.

Participant categories are not monolithic. First, whereas AMPLIFY "members" largely describes East African leaders based in the communities where their organizations operate, in four cases "members" were North Americans collaborating with East African co-leads (3) or working under an East African employer (1). In all four of these cases the East African leader/employer was also invited to the study, and in two instances they also participated. Five member-participants also serve as board members of AMPLIFY Girls.

Donor organizations ranged in size from small family foundations to large private international foundations operating around the world. In all cases, donors were focused on funding efforts to support women and adolescent girls. Two of the donor participants were also East African, the rest North American. Lastly, AMPLIFY staff (three) were North American (one) and East African (two) and located in both geographic locations. Table 2 provides a

breakdown of participants by category of stakeholder, the number invited and the number who actually participated.

Table 2 *Participants by Type*

Participants by Type			
	# invited	# participated	# of organizations
Members	24	19	17
Staff	3	2	1
Donors	10	8	6
Researcher	1	1	N/A
Total	38	30	24

Group Concept Mapping Process

Concept mapping was conducted in five stages for this study (corresponding to stages 2–5 in Kane & Trochim, 2007): 1. Idea generation, 2. Idea synthesis, 3. Idea sorting and rating, 4. Analysis, and 5. Interpretation. Data collection was carried out in person and online between August 2019 and May 2020. Idea sorting and rating and data analysis (stages 3–4) were carried out online using Group Wisdom software.[28] Idea generation and synthesis was conducted in August 2019 in Nairobi in person for members, and via email for all donors but one who also attended the in-person meeting. Interpretation was originally intended to be an in-person workshop, but due to the COVID-19 pandemic this was also conducted online in a two-hour long meeting with AMPLIFY members, in May 2020.

In order to improve the external validity of the results through intentional heterogeneity of ideas and participants (Trochim & Linton, 1986), both AMPLIFY members and donors were included in the first and most important stage: generation of ideas. However, in order to preserve

[28] Group Wisdom software is provided courtesy of Concept Systems Inc.

the privileging of local expertise (Brisolara et al., 2014), AMPLIFY members' feedback led the interpretation stage of the research and their rate-data was the only data used for production of the final concept map solution. Donors were asked to participate in all three stages of data collection, but their rate data was used as a point of comparison only.

AMPLIFY staff (just two people) and myself participated in sorting and rating, but, similar to donors, staff/researcher rate data was used for additional comparative analysis, not to produce the final map. Although members and donors contributed to the field of ideas through brainstorming, members' data made up 87% of the final statement set used in concept mapping. Below is a brief description of the procedure in each stage. A detailed view of who participated in each stage of data collection by constituency is displayed in Table 3.

Table 3 *Participants by Stage of Data Collection*

Participants by Stage of Data Collection			
	Brainstorm	Sorting	Rating
Members	18	15	17
Donors	7	4	5
Staff	not invited	2	2
Researcher	not invited	1	1
Total	25	22	25

Stage 1: Idea generation. Members and donors were asked to brainstorm an exhaustive list of responses to the question *what program components lead to success for girls?* This organizing question was generated collaboratively with member participants, based on a year of evaluation discussions in which "agency," had repeatedly been identified by organizations as a key indicator of girls' holistic success. Response statements were recorded on paper by each participant attending an in-person meeting of the collective in August 2020 and were later input into an excel spreadsheet. Donors not in attendance contributed their statements via email, which were

added to the spreadsheet. Brainstorming generated a total of 406 individual statements from 26 participants (18 members and 6 donors).

Stage 2: Idea Synthesis. Statement data was "cleaned" to ensure that each statement represented a single, unique, response. Statement data was synthesized in five rounds. In the first round, statements were evaluated for repetition only and those statements that were identically repetitive were dropped (this initially reduced the list by approximately 100 statements). Later rounds of synthesis involved thematic coding of each statement by key word and sorting statements based on code, which then allowed statements that were alike to be identified and evaluated for repetition, overlap, or multiple components. Statements that overlapped were re-written to represent single, unique domains, those that contained multiple ideas were separated into two or more statements, and those that were repetitive were dropped. Idea synthesis reduced the total number of statements from 406 to 123.

Stage 3: Idea sorting and rating. Sort-rate participants read each of the final 123 statements and sorted them into piles representing conceptual categories that made sense to the participant. Once sorted, participants were encouraged to provide a "title" for each pile that represented its contents. After sorting, participants rated each statement on a scale of 1–5 (where 1 indicates "not important" and 5 indicates "essential") according to the following prompt: *how important is this component to improving girls' agency?*

Stages 4 and 5: Analysis and Interpretation. Sort data was aggregated into a similarity matrix representing the frequency with which each statement was sorted together with another. This

similarity matrix was analyzed using two-dimensional multidimensional scaling (MDS) to produce a "point map" with each point representing a statement, and its location on the map determined in spatial relativity to how often it was sorted with the other statements (Kane & Trochim, 2007, p.92–93). In other words, two points appearing close together were sorted together more frequently by participants, and therefore are viewed as more conceptually related by participants, whereas points farthest apart from each other on the map were rarely (if at all) sorted into the same pile. A stress value for the point map was calculated to roughly indicate the extent to which the map accurately captured in two-dimensional space the sort data produced.

Next, the point map was grouped into thematic categories using hierarchical cluster analysis specified with Ward's method, which is an agglomerative cluster analysis approach that produces as many clusters as there are points in the data (Kane & Trochim, 2007, p. 99). Thus sense-making of cluster maps really depends on the researcher and participant feedback. I reviewed possible cluster solutions to the map and presented the three most "interpretable" versions of the map to members (6-cluster solution, 7-cluster solution, and 8-cluster solution) at an online meeting.

I determined "interpretability" based on the balance of complexity and simplicity which the map showed for building a theory of practice, and for the coherence of the clusters. As clusters got larger, for example, their statement contents became less thematically coherent. Conversely, as clusters got smaller, they became repetitive, with multiple clusters representing very similar categories of practices. The three Cluster-Map solutions presented to members for discussion represented a middle ground that was neither too repetitive nor too incoherent.

Once a final cluster solution was selected by members, clusters were overlayed with participant rate data to produce a Cluster-Rating Map. Each cluster on the Cluster-Rating Map

has between one and five "layers" that corresponds to the average rating aggregated from all points in the cluster, showing visually each thematic category of practice and its relative importance for improving girls' agency. Several Cluster-Rating Maps were considered, including those with members' rating data only, those with donor rating only, and those with combined ratings. I found substantial discrepancy between members' rating and donors', and so specified the final cluster-rating map with rate data from members only, intentionally privileging practitioner expertise.

All clusters were initially labeled using key words from participants' own titles of conceptual piles during sorting. Words that most frequently appeared in the pile titles of a given cluster, were identified by the Group Wisdom software and put together as labels for each cluster. In some cases these labels clearly described the content of each cluster but in others, the labels were vague and hard to interpret. Thus, cluster labels were reviewed by participants for interpretability and new labels were suggested. I also conducted a thematic analysis of the statements within each cluster, grouping statements into key "practice areas," which contributed to the refinement of three cluster labels to more accurately capture thematic content. Finally, I analyzed the "conceptual fit" of clusters by calculating and reviewing the bridging values for each statement and cluster, and by reviewing the thematic categories of practices that emerged in each cluster after thematic coding.

Results

Cluster-Rating Map: Seven Categories of Practices for Improving Agency

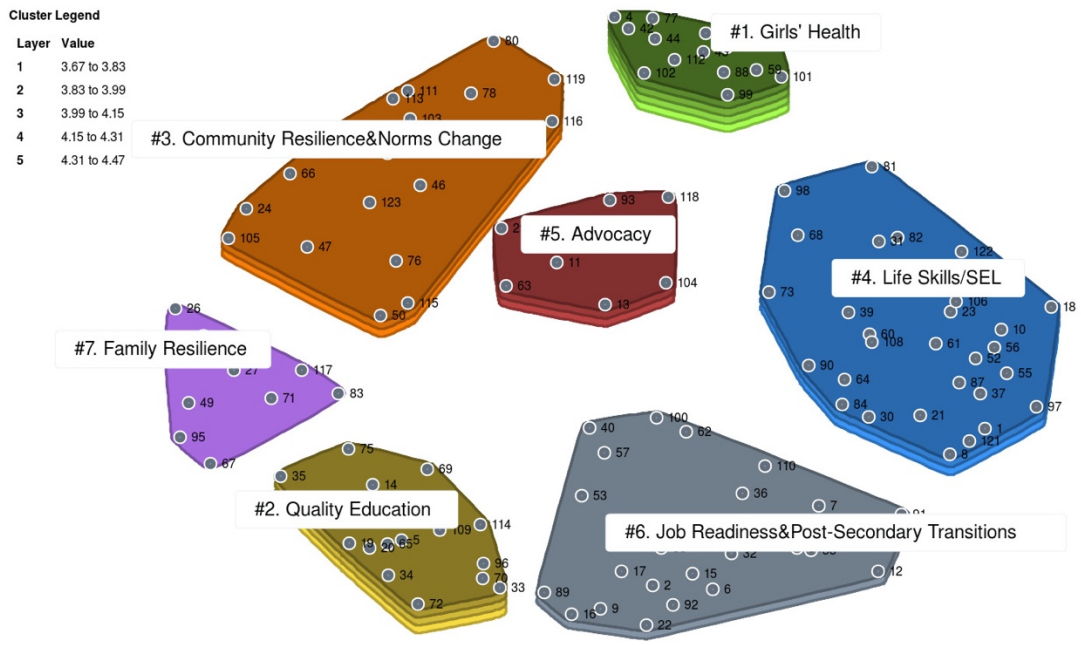


Figure 1: *Final Cluster-Rating Map-Practices that Improve Girls' Agency*

In workshop discussion with member-participants, consensus was quickly reached that the 7-cluster solution of the map best represented the field of practices of the group. The 7 cluster-rating map is displayed in Figure 1. Each numbered point on the map corresponds to a specific statement generated by participants and its location is represented in two-dimensional spatial proximity to those most conceptually related or distant from it. The boundaried clusters and corresponding titles show thematic categories of practice as determined by a combination of hierarchical cluster analysis and qualitative review by participants. Each cluster also has 1-5 layers which represent the average importance rating given to statements in the cluster for girls agency. In other words, statements in clusters with more layers had a higher average importance rating for agency than those with fewer. Overall, the cluster-rating map shows seven thematic categories of practices for improving girls' agency, each with a different "importance" rating.

Standard deviations for each cluster's rating scores were also calculated. Cluster #1—Girls' Health (average rating score 4.47/5, $\sigma = .21$)—followed by cluster #2—Quality Education (average rating 4.15, $\sigma = .32$)—were identified as most important for girls' agency. The next three clusters had near-equal rating and slightly higher standard deviations: Cluster #3: Community Resilience & Norm Change (4.08, $\sigma = .38$). Cluster #4: Life Skills/SEL (Social Emotional Learning) (4.07, $\sigma = .33$), and Cluster #5: Advocacy (4.07, $\sigma = .37$). Cluster #6—Job Readiness and Post-Secondary Transitions—had an average rating of 3.86 ($\sigma = .31$), followed by cluster #7. Family resilience (3.67, $\sigma = .33$).

The associated point map for the 7-cluster solution had a reported stress value of .3197. Kane & Trochim suggest a range of .205-.365 as "common" for 95% of Concept Mapping exercises (2007, p. 98). The stress value can be interpreted as a rough indicator of how well the map represents the complexity of the data in a bivariate plot. A stress value of .3197 is on the higher end of the range, indicating that there are likely some areas of the map where there is less consensus on conceptual relationships amongst the participants and where the MDS operation did not precisely represent participant data. Some amount of "error" in the visual representation of conceptual categories is to be expected given the diversity of programs, participants, and geographic locations represented in the study, and can be better understood by examining the conceptual fit of each cluster and its associated statements.

In order to dig deeper on the question of "accurate" representation, I examined bridging values and conducted a qualitative thematic coding of practices in each cluster. Bridging values are a rough indicator of whether a statement was placed in its current location on the map because of the high frequency with which it was sorted with the statements immediately surrounding it or whether it was placed in an intermediate position because it was sorted at a

low-to-medium frequency with the more distant statements on either side of it (Kane & Trochim, 2007). Thus, clusters with a low average bridging value usually represent a more conceptually bounded cluster than the others. The specific practice categories making up each cluster (identified through qualitative coding), the associated bridging values and rating of each statement and cluster can be viewed in Table 4 (in the appendix). In clusters with higher average bridging values and less thematic clarity there are typically one or two statements that seem "misplaced." These are noted with an (**) in Table 4.

At first glance, the bridging values and qualitative analysis seem to indicate that the Life Skills/SEL cluster (average bridging value of .28), Girls' Health (average bridging value of .13), and Job Readiness and Post-Secondary Transitions (average bridging value of .26) are tightly-bounded conceptual categories of practices. Girls' Health was originally named "Improving SRH outcomes and understanding," a title derived from participants' pile titles during sorting. After qualitative coding of practice types, I simplified this title to "Girls' Health" because the statements contained in the cluster related to either direct provision of health services to girls, or health education (usually in the area of sexual and reproductive health and gender-based violence awareness and prevention). In this cluster there is no single statement with a bridging value above .22 and all the statements have a clear relationship to promoting and maintaining "girls' health."

Similar conceptual clarity can be found in the Life Skills/SEL and Job Readiness and Post-Secondary Transitions cluster, where I identified four distinct types of practices in each cluster which were all closely related to the cluster title: soft skills training, opportunities to practice soft skills, provision of mentors or mentoring relationships, and peer communities of

support. Similarly, the Job Readiness cluster included four practice categories: job skills training, vocational training, exposure to careers, and post-secondary transitions support.

Community Resilience & Norms Change both had a high average bridging value (.45) and a relatively high average rating (4.08 or "very important," standard deviation .38). After thematic coding of statements, the types of practices emerging from this cluster make better theoretical sense, involving activities and initiatives that would make communities "better able to support girls' needs, and rights." Herein we see three primary categories of practices: 1. Support for girls' needs (physical needs as well as social emotional needs), 2. Sensitization for communities, girls, and boys on harmful gender norms, gender-based violence and girls' rights, and 3. Legal support protecting girls' rights. Implicit here is the idea that attending to normative change goes hand in hand with providing girls the resources they need to lead safe, healthy lives. I interpret high bridging values in this cluster as indicative of the fact that the target group of suggested interventions are diverse (including girls themselves, boys and men, community leaders and parents) as are the types of needs addressed (basic needs, legal rights, awareness, etc.).

Advocacy occupies both the physical middle ground of the cluster map as well as the theoretical middle ground bridging Soft Skills/SEL to Community/Norm Change. After qualitative coding, I found the statements in the Advocacy Cluster to be related to two different types of "opportunities" for girls: opportunities for girls to engage with their communities (school and home) and opportunities for them to learn about and celebrate themselves and their rights in a safe setting. In later iterations of this map, I changed the name from "Advocacy" to "Opportunities to Engage," based on the thematic codes.

Although it is straightforward in title, Cluster #2, Quality Education, takes a holistic view of education. Thematic coding produced four categories of practices: 1. initiatives to improve school quality through infrastructure and better, more diverse, academic offerings, 2. programs to make schools more accessible to girls through learning materials and scholarships, 3. improved pedagogy and 4. teacher training. Notably, a number of practices grouped here refer to training teachers to utilize a better pedagogy—one that is responsive to gender and is student-centered. The original title of this cluster given by partners was "Better Schools, Better Pedagogy," which was changed to "Quality Education" at the virtual workshop by AMPLIFY member-participants.

Interpreting Cluster-Ratings

While there is some variation of rating scores between clusters, the lowest average cluster rating is 3.67 (Family Resilience) and the highest is 4.47 (Girls Health). Both the highest and the lowest scores indicate that the practices were viewed by participants as important for agency, where a score of "3" denotes "somewhat important," a score of "4" indicates "very important," and a score of "5" indicates "essential." One way to interpret this small range of rate values is that all practice categories represented on the map are important for girls' agency, but that no single area of practice is "essential" on its own, suggesting a holistic approach to girls' agency. However, given that these ratings did not use a forced ranking system, it is hard to find this interpretation conclusive. To get a sense of the extent to which these ratings represent group consensus, I examined the relative rank of each cluster (as indicated by average cluster ratings) by country and by participant category.

Table 5 shows a quick comparison of relative rank of clusters and average cluster ratings across categories of participants. Column 1 shows the average cluster rank (and corresponding average rating) of the final concept map in Figure 1, while columns 2–5 show a country-level

breakdown of relative rankings. Columns 6 and 7 show data for staff and the researcher, as well as donors, that were not included in the final cluster-rating solution.

Table 5 *Average Cluster Rank by Participant Category*

Average Cluster Rating and Cluster Rank by Participant Category							
	Cluster Map	Tanzania	Kenya	Rwanda	Uganda	Staff & Researcher	Donors
Cluster Name	n=17	n=4	n=6	n=4	n=3	n=3	n=5
Girls Health	#1 (4.47)	#1 (4.29)	#1 (4.59)	#1 (4.75)	#1 (4.14)	#1 (4.12)	#2 (3.88)
Quality Education	#2 (4.15)	#2 (4.15)	#5 (4.16)	#2 (4.51)	#4 (3.69)	#4 (3.76)	#4 (3.77)
Community Resilience/Norms	#3 (4.08)	#5 (3.87)	#3 (4.19)	#3 (4.43)	#2 (3.73)	#3 (3.78)	#3 (3.83)
Life Skills/SEL	#4 (4.07)	#4 (3.89)	#4 (4.17)	#4 (4.39)	#3 (3.71)	#2 (3.90)	#1 (3.93)
Advocacy	#5 (4.07)	#3 (4.04)	#2 (4.24)	#6 (4.15)	#5 (3.67)	#5 (3.67)	#5 (3.77)
Job Readiness & Post-Secondary Transitions	#6 (3.86)	#6 (3.78)	#6 (3.87)	#5 (4.22)	#6 (3.50)	#6 (3.44)	#7 (3.39)
Family Resilience	#7 (3.67)	#7 (3.41)	#7 (3.79)	#7 (3.90)	#7 (3.48)	#7 (2.88)	#6 (3.40)

The table shows consensus amongst participants from different countries on the relative importance of Girls' Health (#1) and Family Resilience (#7), and near-consensus on Life Skills/SEL in position #4 and Job Readiness & Post-Secondary Transitions in position #6. However, relative rankings vary a good deal between countries for three clusters (Quality Education, Advocacy, Community Resilience and Norms) indicating that the importance of these practice categories vary based on locale. Staff/researcher ratings were very similar to AMPLIFY participants' overall ratings, while Donor ratings were very different—producing a #1 rank for Life Skills/SEL, which held 4th position for AMPLIFY participants.

Triangulating Findings with the Literature

The findings presented herein confirm and expand on what the literature communicates about best practices for improving girls' agency. Practices gleaned from the literature review (Table 1) are found represented in the thematic practice categories of the final Concept Map. The most frequently-appearing literature-identified practice is student-centered pedagogy, which is located in at least three of the seven clusters of practices on our Concept Map, including Quality Education (where "student centered pedagogy" appears as a stand-alone statement), Girls' Health (where peer mentoring and girl-led trainings in SRH were identified in 3 statements), and of course in Life Skills/SEL where peer-led curriculum and opportunities for girls to practice soft skills reflect a student-centered ethos.

Structured and facilitated soft skills curriculum is articulated clearly in the Life Skills/SEL cluster, as are female role models and peer communities of support, which were also identified as thematic categories of practice in the Soft Skills cluster. Although less obviously featured, practices that might include critical pedagogy appeared in Girls' Health in statements such as "SRH programs for girls that address gender and power," and in the Quality Education cluster in statements calling for alternative teacher training or alternative pedagogical approaches.

The principal of holistic care is represented clearly in the Concept Map, where clusters on the left side represent one of the three social spheres (community, family, and school) and within each of these three clusters multiple stakeholders are identified, including parents, teachers, boys/men, and community leaders. On the right side of the map, members seem to articulate a vision for providing care for multiple aspects of girls' well-being in clusters addressing girls' physical well-being, social-emotional well-being, and financial well-being (Girls' Health, Life Skills/SEL and Job Readiness).

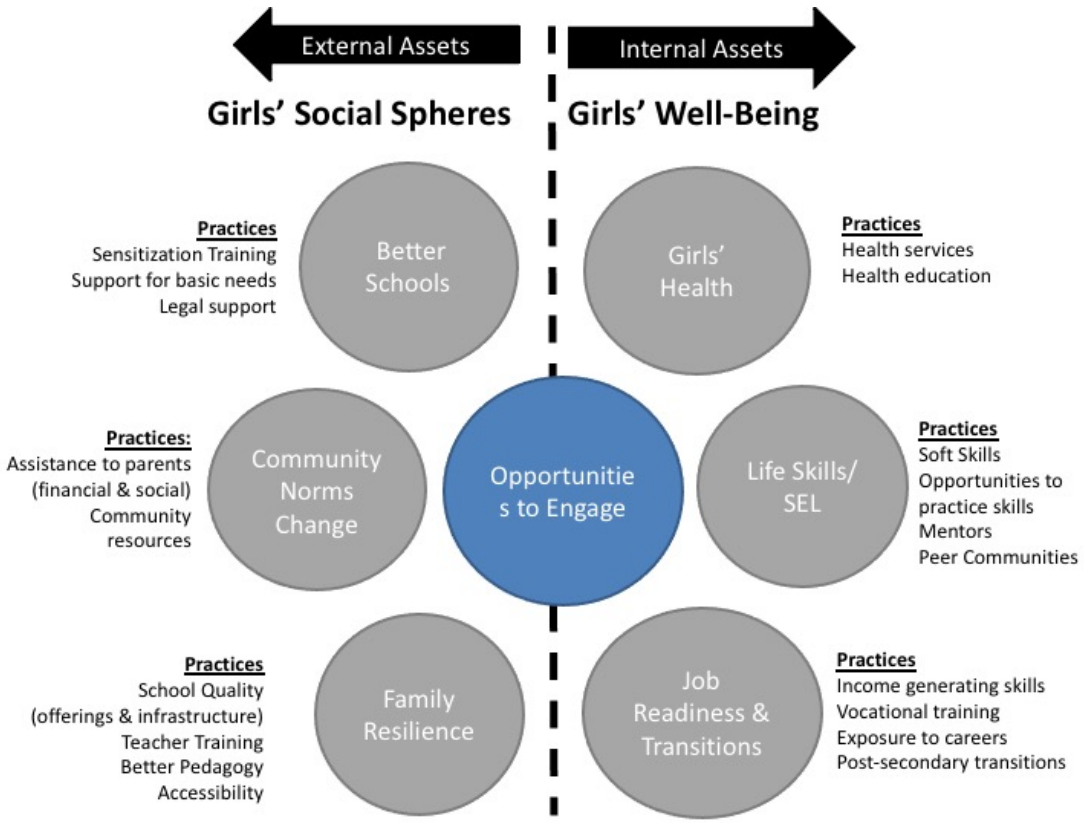
Such holism echoes the literature and the view of youth development as a "process" (Bronfenbrenner & Morris, 1994) engaging multiple stakeholders, multiple social spheres, and attending to multiple facets of girls' well-being (Banati et al., 2021; Temin & Heck, 2020; Murphy-Graham and Leal, 2015; and Catalano et al., 2004). In a review of 94 positive youth development (PYD) programs in LMICs, Alvarado et al. found that most PYD programs engaged community members and youth and were implemented in multiple social settings (2017). This echoed Catalano et al.'s study of PYD programs in the US, which found that programs operating in multiple social domains were able to address a broader range of youth outcomes (Catalano et al., 2004).

The divide between social spheres on the left and individual well-being on the right mirrors Benson's distinction of "external" and "internal" assets (2003, 2011). Unlike Benson, however, where internal assets are characterized by the traits, attitudes, values, and competencies that support adolescent well-being and external assets cross multiple social spheres (2011, p. 199–200), AMPLIFY's framework ties internal assets to specific spheres of well-being (health, financial, social, emotional), and external assets are linked to the specific social contexts of community, school, and family. On this view, advocacy can be interpreted as a "bridging category"—one that links internal to external assets through a set of opportunities for a girl to engage meaningfully with the different social spheres while exercising the skills and capacities she has developed from the practices attending to her well-being.

Discussion

Synthesizing Learning into Theory

Figure 2: *Draft Theory of Practice for Improving Girls' Agency*



In Figure 2 I translate the Cluster-Rating Map (see Figure 1) into a community-driven draft Theory of Practice for Improving Girls' Agency. I do this by standardizing the size of each cluster and dividing the map into two sections—i.e. clusters that focus on internal versus external assets (Benson, 2003). I then re-title each cluster based on either the social sphere (community, family, or school) or the realm of well-being (health, financial, or social-emotional) that is the primary target of the practices in each category. Next to each sphere is the list of extracted practice categories in each cluster that address the needs of girls in each setting or in each area of wellness (derived from thematic coding).

Social spheres and realms of well-being more or less correspond to the external and internal assets proposed by youth development practitioners (Benson, 2003; Damon & Gregory, 2003). But instead of "infrastructures," our map proposes three areas of necessary engagement with girls' social settings. Within each social sphere on the map there are practices that are individually targeted (that is, they support a girl in the specified setting) or that target key stakeholders in a girl's community by improving the setting through engagement with parents, teachers, community leaders, and school buildings, etc. Herein we see an explicit acknowledgement that success for girls means they must have the tools to navigate their social settings and that social settings must shift to be more supportive of a girl's needs and rights. Social spheres and well-being clusters are tied together by the middle cluster, "opportunities to engage," in which girls are given the opportunity to practice what they've learned from programming in an area of well-being and to deploy this in one of their social settings.

Importance of Girls' Health

Our Concept Map also shows some important departures from, and expansion on, the literature, notably the emphasis given to Girls' Health in terms of both services and health education. While a few studies evaluate the impact of health services in connection to a life skills curriculum (e.g. Austrian et al., 2016), most studies in our review consider health education to be part of, rather than distinct from, soft skills programs. This conflation of two distinct content areas has potentially contributed to the under-valuing of the importance of girls' health to agency and the lack of consensus on life-skills content more broadly in the sector (Murphy-Graham & Lloyd, 2016). Often, we find girls' education initiatives tackling either health or quality education, and only informal programs such as life skills curricula bridge the divide between the two.

Community experts clearly articulate that girls' health should be a top priority of programs addressing their agency, and that it should go hand in hand with quality education.

Making Social Settings More Supportive for Girls and Addressing Her Holistic Well-Being

Secondly, approaches to improving girls' agency must ensure that the primary social settings in which girls' lives unfold are characterized by some basic understanding and support for girls' rights and needs. This includes legal supports, material support, and community, school and family settings that are engaged in the work of deconstructing harmful gender norms and of making these spaces more supportive of girls.

Lastly, replicating success means preparing girls with the skills and knowledge that they need to develop agency-related skills (self-governance, leadership, self-beliefs, etc.), but also with skills related to two other aspects of their well-being: health and financial independence. Program practice has long understood the former point, but it is only recently beginning to acknowledge that the necessary skills for a girl's agentic development must address her physical and financial health in addition to her social-emotional well-being. Skills-based learning should go hand in hand with safe opportunities to engage in girl-led activities where girls are able to use their skills for broader engagement with their communities and families.

Contribution of the Method & Theory for Driving Practice

I began this paper by arguing that the field of practice requires replicable context-inclusive theories rather than replicable program interventions. Whereas experimental research largely seeks to isolate a set of factors (e.g. program interventions) and prove their impact, both recent evidence (Hanushek, 2021; Buehren, 2017; Klocker, 2016; Murphy Graham & Leal, 2015) and logic suggest that those factors exist within a broader environment which influences their effects.

Thus experimental evaluations must be combined with methods that increase external validity and, specifically, cross-country or cross-population generalizability.

Group Concept Mapping as a theory-building method offers several advantages to experimental methods for identification of best practices. Firstly, there is the opportunity to sample for heterogeneity and approach generalizability using alternative approaches to sampling (Cook & Campbell, 1979), since the complexity of securing a representative sampling from an international population poses serious practical problems. The heterogeneity represented by our sample population arguably makes a better case for the generalizability of our theory of practice to adolescent girl populations growing up in resource-limited settings in study countries than RCT results from a single country.

Second, the method is participatory, and offers a way to center local practitioner expertise and Southern leaders in practice conversations about how to serve their communities. This results in more robust and localized conceptualizations of practice, while undermining traditional power hierarchies in terms of who is "qualified to know" (Brisolara 2014). Finally, Group Concept Mapping is designed as a conceptualization method, which makes it a valuable tool for theorizing "categories" of practices as opposed to specific interventions. In this way it is a more natural method for developing an "ecosystems" approach to questions of youth development.

Limitations

Our exercise did not, however, provide specific or conclusive information about how certain structural constraints vary country to country, but rather provided clues in terms of variable rate-data that such variation did in fact exist. Other methods might better support the identification of discrete environmental constraints through contextual analysis, such as social decomposition, which examines the variability of phenomenon across a population to identify the underlying drivers of social change (Eloundou-Enyegue, Giroux, & Tenikue, 2021), or multi-level

modeling, which supports the identification of group-level or community level "effects" on outcomes at the macro level. These alternatives, however, are not suitable for producing a holistic view of context, but rather for identifying singular or macro-level contextual contributors.

Group Concept Mapping, as I've deployed it in this study, also doesn't tell us much about within-country variation, although using a different sampling strategy and a larger sample size within each country does give Group Concept Mapping the potential to provide some of this insight. Furthermore, while my analysis theorizes what categories of practice are important for girls' agency, it does not explain how or why these categories matter—in other words, it does not offer a theory of causal mechanism or causal complexity. For these explanations we require qualitative inquiry with girls and implementers about how programs work and how they are affected by contextual factors. In fact, I argue that qualitative investigations into causal complexity would be a natural extension of this research and would help fill in the theoretical picture of girls' agency program practice.

Although this study centered the *needs* of adolescent girls in the inquiry (Brisolara, 2014), I fail to include girls' own voices, instead trusting the practitioners who work with girls to represent their needs. This is a clear limitation of the study. More research is needed to understand and clearly articulate the specific aspects of context in each social sphere that matter for girls' agency and to gain a more detailed view of how these interact with each other and with girls' well-being. Girls themselves are best positioned to offer this detail, and I recommend direct engagement with girls as a priority for further study. Participants' views on the subject of best practice would help to strengthen and refine the theory proposed through triangulation of knowledge with practitioners.

Conclusion: Theory as a Heuristic for Replicating Outcomes

In this paper I offer a draft theory for improving girls' agency as a heuristic tool that can be used by program practitioners to guide program design for replicating girls' agency outcomes, and for testing the effects of these program designs locally and/or regionally. The theory suggests that programs which engage with multiple social spheres (school, community, and family) and attend to multiple facets of girls' well-being (e.g. financial, health, psychosocial) will increase agency outcomes. Girls' Health is a top priority for agency, although the relative importance of other programmatic areas is less clear and likely variable based on context. Accordingly, which social spheres and aspects of well-being yield the biggest effect, and, further, which program practices within each sphere are most important and in which settings, are questions that require further research and localized testing.

I urge further research that would test this theory's internal validity for multiple geographies, and qualitative inquiry with girls, that might shed light on how these categories of practices interact with each other to produce agency. Such investigation of causal complexity will necessitate a wider range of methodological tools beyond experimental studies, and a conceptual shift away from scalable one-size-fits-all programs. Practitioners and policy makers should pair the more traditional tools for producing internally valid "evidence" with methods that build a stronger and more explicit case for cross-country generalizability, and which offer opportunities for the contribution of local expertise. Girls' specific, gendered, and contextualized needs require that we bring local thinking and practitioner-driven theory to bear on the question of "best practice."

Table 4 Clusters Statements with Practice Area, Rating and Bridging Values

Statement #	Statements	Practice Area	Rating	Bridging Value
	Cluster name: Improving SRH outcomes and understanding.		**Avg 4.47**	Avg 0.13
4	access to healthcare services	health services	4.82	0.1
25	Health and sanitation training for girls	health training	4.69	0.11
41	access to youth-friendly family planning and SRH services	health services	4.53	0
42	menstrual hygiene supplies provision	health services	4.73	0.14
43	education about responsible/safe sex	health training	4.47	0.07
44	psychological support-counselling & other mental health services	health services	4.35	0.22
59	menstrual health training	health training	4.65	0.22
77	HIV/AIDS testing and counselling	health services	4.38	0.08
79	SRH training for girls including family planning	health training	4.56	0.11
88	safe forums to discuss difficult topics and to debunk stereotypes about sex, health and reproduction.	health training	4.29	0.23
99	girl-led SRH trainings	health training	4.19	0.19
101	peer to peer mentoring in SRH	health training	4.13	0.17
102	programs that sensitize boys on SRH issues	health training for boys	4.29	0.16
112	SRH programs for girls that address gender & power	health training	4.53	0.1
	Cluster name: Better Schools, better pedagogy		**Avg 4.15**	Avg 0.32
5	access to computers or laptops	school quality (offerings)	4.25	0.28
14	guidance and counseling to understand options beyond high school	school quality (offerings)	4.35	0.74
19	scholarships to attend school	accessibility	4.82	0.25
20	provision of learning and school materials (e.g. textbooks, notebooks, pens, school uniforms)	accessibility	4.18	0.22
33	academic counseling (Tracking progress, and identifying student challenges)	school quality (offerings)	4.29	0.2
34	monetary support to schools (e.g. hiring teachers)	school quality (infrastructure)	3.41	0.27
35	parent engagement in student academic progress	school quality	4.33	0.47
65	teacher training to work with girls with disabilities	teacher training	4	0.45
69	income generating opportunities for girls	school quality (offerings)	4.06	0.4
70	academic tutoring	school quality (offerings)	4.19	0.21
72	low student-teacher ratio	school quality (infrastructure)	3.59	0.13
75	access to the internet	school quality (infrastructure)	4.24	0.45
96	student-centered pedagogy	better pedagogy	4.21	0.14

109	teacher training that explicitly discusses dissects, and seeks to change gender social norms	teacher training	4.18	0.33
114	girl-friendly curriculum (gender responsive pedagogy)	better pedagogy	4.18	0.29
	Cluster name: Life Skills/ SEL Training in Schools		**Avg 4.07**	Avg 0.28
1	self-awareness curriculum	soft skills training	4.24	0.28
8	sports for teaching life skills (confidence, conflict resolution, etc.)	soft skills training	3.69	0.34
10	public speaking skills development	soft skills training	4.19	0.28
18	conflict resolution training	soft skills training	3.82	0.36
21	leadership trainings	soft skills training	4.41	0.14
23	self-acceptance and confidence building programming	soft skills training	4.76	0.33
30	training for girls in political campaigning	soft skills training	3.12	0.17
31	political advocacy & lobbying training for girls	soft skills training	3.53	0.43
37	goal setting and goal planning training	soft skills training	4.18	0.19
39	training in empathy and the value of giving back (volunteerism)	soft skills training	4.06	0.25
45	opportunities to engage in personal story telling	opportunities to practice skills	3.82	0.35
52	personal effectiveness and life planning training	soft skills training	4.06	0.15
55	training in effective communication skills	soft skills training	4.38	0.2
56	negotiation skills training	soft skills training	4.06	0.19
60	exposure to inspirational speakers	role models	3.88	0.19
61	opportunities to speak publicly and show confidence	opportunities to practice skills	4.13	0.16
64	opportunities to mentor others	opportunities to practice skills	4.07	0.11
68	peer to peer communities of support	peer communities	4	0.37
73	alumnae role models for girls	mentors/role models	4	0.22
81	sports for physical wellness programs	health**	3.65	0.66
82	peer to peer mentoring in psychosocial support	peer communities	4.12	0.44
84	opportunities to take on leadership roles	opportunities to practice skills	4.44	0.14
87	programming to promote critical thinking and problem solving skills development	soft skills training	4.44	0.23
90	opportunities to be mentored by educated women	mentors/role models	4.5	0.14
97	program that focus on resilience	soft skills training	4.24	0.35
98	self-defense training	soft skills training	3.88	0.5
106	performing arts for building positive self-beliefs (self-esteem, confidence, self-efficacy)	soft skills training	4.06	0.34
108	exposure to women and women role-models in non-traditional roles	mentors/role models	4.38	0.18
121	opportunities to practice good decision-making	opportunities to practice skills	4.25	0.27
122	mindfulness and well-being training	soft skills training	3.94	0.49
	Cluster name: Advocacy		**Avg 4.07**	Avg 0.34

11	community service opportunities	opportunities to engage	3.76	0.32
13	opportunities for girls to teach in their schools and communities	opportunities to engage	4.06	0.3
29	parents training in psychosocial support for their girls	parent training**	3.81	0.44
63	safe space for girls	opportunities for girls to learn about themselves and rights	4.53	0.34
93	events to celebrate girl-hood	opportunities to engage	3.65	0.32
104	programs that engage girls during school breaks	opportunities to engage	4.06	0.31
118	teach girls their rights and laws that protect them	opportunities for girls to learn about themselves and rights	4.59	0.35
	Cluster name: Community/Norms Change		**Avg 4.08**	Avg 0.45
24	child care provision for teen mothers	support for pregnant girls	3.69	0.46
46	shelter/accommodation, clothing and other basic needs for vulnerable students	support for girls' basic needs	4.31	0.37
47	family mediation services	support for families	3	0.62
50	household visits	support for families	3.53	0.46
51	advocacy on behalf of girls for their rights protection	legal support	4.47	0.57
66	training/sensitization with local leaders on girls' issues	sensitization for communities	4.24	0.53
76	programs that ensure girls access to food & healthy nutrition (school feeding programs, etc.)	support for girls' basic needs	4.19	0.53
78	clean water provision for girls' households	support for families' basic needs	4.31	0.43
80	mosquito net provision for malaria prevention	support for girls' basic needs	3.69	0.44
94	programs in communities that explicitly address harmful gender norms	sensitization for communities	4.35	0.41
103	safe space for girls to live during school breaks	support for girls' basic needs	3.73	0.36
105	community center	support for families and communities	3.56	0.6
107	performing arts for delivering important information re. SRH, GBV and gender norms	sensitization for communities	3.75	0.44
111	GBV awareness and prevention programs for girls that explicitly addresses gender and power	sensitization for girls	4.53	0.51
113	programs for men and boys that explicitly address harmful gender norms & GBV prevention	sensitization for communities	4.35	0.36
115	policies protecting girls from violence in school	legal support for girls	4.69	0.41
116	policies de-stigmatizing sharing reproductive information at school	legal support	4.29	0.39
119	access to services for GBV	support for girls basic needs	4.41	0.26

120	programs that build awareness in communities for supporting girls' health, growth, education and well-being	sensitization for communities	4.38	0.43
123	girl-led advocacy campaigns in communities	sensitization for communities	4.06	0.47
	Cluster name: Job readiness - Transitions	**Avg 3.72**	**Avg 3.86**	Avg 0.26
2	curriculum in tech skills & computer science	job skills training	3.71	0.26
3	internships or apprenticeships in industry	vocational training opportunities	3.82	0.17
6	hands-on vocational training in various industries (including STEM fields)	vocational training opportunities	3.94	0.17
7	entrepreneurship training for girls	job skills training	4.24	0.29
9	fine arts skills development	job skills training	3.19	0.34
12	exposure to international exchange/interaction	exposure	3.24	0.39
15	student exposure tours to universities, places of employment, and historic sites	exposure	3.88	0.24
16	english fluency classes	job skills training	3.81	0.29
17	computer skills training	job skills training	4.24	0.24
22	coaching girls through university applications	transitions support	4.13	0.31
32	career exposure--career fair, career visits	exposure	4.06	0.15
36	teacher training in mentoring and psychosocial support to students	transitions support	3.94	0.38
38	training on labor market and different skills required for different jobs	job skills training	3.94	0.23
40	opportunities to engage in and execute entrepreneurship	job skills training	4.18	0.24
53	support identifying and applying for jobs	transitions support	4.13	0.28
54	employability skills training	job skills training	4.25	0.28
57	savings, budgeting and money management training for girls	job skills training	3.87	0.29
58	artisan skills training-making things to sell/reusable pads	job skills training	3.41	0.4
62	consistency of engagement (weekly meetings)		3.82	0.29
85	training for girls in agriculture and gardening	job skills training	3.41	0.18
86	opportunities to engage in local and regional competitions in areas of interest for girls (e.g. STEM, debate, etc.)	exposure	3.59	0.18
89	extracurricular clubs in specific areas-math, debate, engineering	exposure	3.94	0.2
91	program planning skills building	job skills training	3.71	0.32
92	study skills training	transitions support	4	0.29
100	programs that engage girls productively during gap-year	transitions support	3.76	0.18
110	experiential/hands-on and project-based learning	job skills training	4.29	0.19
	Cluster name: Family Resilience		**Avg 3.67**	Avg 0.57
26	conflict management training for parents	assistance to parents	3.65	0.65
27	savings training for parents	$ assistance to parents	3.76	0.52
28	training parents in income generating projects	$ assistance to parents	3.82	0.6
48	emergency financial assistance to vulnerable families	$ assistance to parents	3.88	0.48
49	community savings and loans groups for parents	$ assistance to parents	3.47	0.6

67	microloans	$ assistance to parents	3.13	0.61
71	free community library	community resources	3.94	0.46
74	provision of electricity	community resources	3.59	0.44
83	school policy to retain students who are pregnant	policy	4.31	0.48
95	local staff/hiring from within	community resources	3.47	1
117	programs to ensure girls' safe transport to school (e.g. provision of bicycles)	community resources	3.31	0.47

References

Alvarado, G., Kapungu, C., Moss, C., Plaut, D., Reavley, N., & Skinner, M. (2017). *A Systematic Review of Positive Youth Development Programs in Low-and Middle-Income Countries*. YouthPower Learning and Making Cents International. https://pdf.usaid.gov/pdf_docs/PA00MR58.pdf

Amin, S., Ahmed, J., Saha, J., Hossain, M., & Haque, E. (2016). Delaying child marriage through community-based skills-development programs for girls: Results from a randomized controlled study in rural Bangladesh. (Report). The Population Council.

Archibald, T., Guisse, B. K., Nidaye, A., Kane, F., & Diouf, F. (2021). Positive Youth Development in Senegal: A case study of 4-H Senegal. *Journal of Youth Development*, *16*(2–3), 344–362.

Austrian, K., Hewett, P. C., Soler-Hampejsek, E., Bozzani, F., Behrman, J. R., & Digitale, J. (2016). *Adolescent Girls Empowerment Programme: Research and evaluation mid-term technical report*. https://knowledgecommons.popcouncil.org/cgi/viewcontent.cgi?article=1560&context=departments_sbsr-pgy

Bajaj, M. (2015). "Pedagogies of resistance" and critical peace education praxis. *Journal of Peace Education*, *12*(2), 154–166.

Baldridge, B. J., Beck, N., Medina, J. C., & Reeves, M. A. (2017). Toward a new understanding of community-based education: The role of community-based educational spaces in disrupting inequality for minoritized youth. *Review of Research in Education*, *41*(1), 381–402.

Banati, P., Rumble, L., Jones, N., & Hendriks, S. (2021). Agency and Empowerment for Adolescent Girls: An Intentional Approach to Policy and Programming. *Journal of Youth Development*, *16*(2–3), 239–254.

Bronfenbrenner, U., & Morris, P. A. (1994). Ecological Models of Human Development. In M. Gauvain & M. Cole (Eds.), *Readings on the Development of Children* (Vol. 1, pp. 37–43). Worth Publishers.

Benson, P. L., Scales, P. C., & Syvertsen, A. K. (2011). The contribution of the developmental assets framework to positive youth development theory and practice. *Advances in Child Development and Behavior*, *41*, 197–230.

Benson, P. L. (2003). Developmental assets and asset-building community: Conceptual and empirical foundations. In R. M. Lerner & P. L. Benson (Eds.), *Developmental Assets and Asset-Building Communities: Implications for Research, Policy and Practice* (pp. 19–43). Springer.

Biesta, G. (2007). Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory*, *57*(1), 1–22.

Brisolara, S., Seigart, D., & SenGupta, S. (2014). *Feminist evaluation and research: Theory and practice*. Guilford Publications.

Buehren, N., Goldstein, M., Gulesci, S., Sulaiman, M., & Yam, V. (2017). Evaluation of an adolescent development program for girls in Tanzania. *World Bank Policy Research Working Paper*, *7961*.

Campbell, C., & Mannell, J. (2016). Conceptualising the agency of highly marginalised women: Intimate partner violence in extreme settings. *Global Public Health*, *11*(1–2), 1–16.

Casey, K. (2018). Radical decentralization: Does community-driven development work? *Annual Review of Economics*, *10*, 139–163.

Catalano, R. F., Skinner, M. L., Alvarado, G., Kapungu, C., Reavley, N., Patton, G. C., Jessee, C., Plaut, D., Moss, C., & Bennett, K. (2019). Positive youth development programs in low- and middle-income countries: A conceptual framework and systematic review of efficacy. *Journal of Adolescent Health*, *65*(1), 15–31.

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (2004). Positive Youth Development in the United States: Research Findings on Evaluations of Positive Youth Development Programs. *Annals of the American Academy of Political and Social Science*, *591*, 98–124. https://doi.org/10.1177/0002716203260102.

Chávez, K. R., & Griffin, C. L. (2009). Power, feminisms, and coalitional agency: Inviting and enacting difficult dialogues. *Women's Studies in Communication*, *32*(1), 1–11.

Conn, K. M. (2017). Identifying Effective Education Interventions in Sub-Saharan Africa: A Meta-Analysis of Impact Evaluations. *Review of Educational Research*, *87*(5), 863–898. https://doi.org/10.3102/0034654317712025.

Cook, T. D., & Campbell, D. T. (1979). Validity. In T.D Cook & D.T. Campbell (Eds.), *Quasi-Experimentation: Design & Analysis Issues for Field Settings* (pp. 37–94). Houghton Mifflin Company.

DeJaeghere, J., & Lee, S. (2011). What matters for marginalized girls and boys: A capabilities approach to exploring marginalization and empowerment in Bangladesh. *Research in Comparative and International Education*, *6*(1), 27–42.

Damon, W., & Gregory, A. (2003). Bringing in a new era in the field of youth development. In R. M. Lerner & P. L. Benson (Eds.), *Developmental assets and asset-building communities* (pp. 47–64). Springer.

Dierker, B. (2016). "You Are Building on Something": Exploring Agency and Belonging Among African American Young Adults. In J. G. DeJaeghere, J. Josic, & K. McClearly (Eds.), *Education and Youth Agency* (pp. 27–46). Springer.

Donald, A., Koolwal, G., Annan, J., Falb, K., & Goldstein, M. (2020). Measuring women's agency. *Feminist Economics*, *26*(3), 200–226.

Eloundou-Enyegue, P., Giroux, S., & Tenikue, M. (2021). Demographic Analysis and the Decomposition of Social Change. In A. Klimczuk (Ed.), *Demographic Analysis-Selected Concepts, Tools, and Applications*. IntechOpen.

Eloundou-Enyegue, P. M., & Stokes, C. S. (2004). Teen fertility and gender inequality in education: A contextual hypothesis. *Demographic Research*, *11*, 305–334.

Haberland, N. (2015). The Case for Addressing Gender & Power. *International Perspectives on Sexual and Reproductive Health*, *41*(1), 31–42. https://doi.org/10.1363/4103115

Hanushek, E. A. (2021). Addressing cross-national generalizability in educational impact evaluation. *International Journal of Educational Development*, *80*, 102318. https://doi.org/10.1016/j.ijedudev.2020.102318

Hull, G., Jury, M., & Sahni, U. (2015). "Son Enough:" Developing Girls' Agency in India across Domestic and Civic Spheres through Feminist Media Practice. In B. Kirshner & E. Middaugh (Eds.), *# youthaction: Becoming Political in the Digital Age* (p. 81). Information Age Publishing.

Kane, M., & Trochim, W. M. K. (2007). *Concept Mapping for Planning and Evaluation* (Vol. 50). Sage Publications, Inc.

Kennelly, J. J. (2009). Youth cultures, activism and agency: Revisiting feminist debates. *Gender and Education*, *21*(3), 259–272.

Kwauk, C., & Braga, A. (2017). *Translating Competencies to Empowered Action: A Framework for Linking Girls' Life Skills to Social Change*. The Brookings Institution.

Kwon, S. A. (2006). Youth of color organizing for juvenile justice. *Beyond Resistance*, 215–228.

Lerner, Richard M., Jacqueline V. Lerner, Jason B. Almerigi, Christina Theokas, Erin Phelps, Steinunn Gestsdottir, Sophie Naudeau, Helena Jelicic, Amy Alberts, Lang Ma, Lisa M. Smith, Deborah L. Bobek, David Richman-Raphael, Isla Simpson, Elise DiDenti Christiansen, & Alexander von Eye. (2005). Positive Youth Development, Participation in Community Youth Development Programs, and Community Contributions of Fifth-Grade Adolescents: Findings From the First Wave Of the 4-H Study of Positive Youth Development. *The Journal of Early Adolescence*, *25*(1), 17–71. https://doi.org/10.1177/0272431604272461

Leventhal, K. S., DeMaria, L. M., Gillham, J. E., Andrew, G., Peabody, J., & Leventhal, S. M. (2016). A psychosocial resilience curriculum provides the "missing piece" to boost adolescent physical health: A randomized control trial of Girls First in India. *Social Science & Medicine*, *161*, 37–46.

Marcus, R., Gupta-Archer, N., Darcy, M., & Page, E. (2017). *Girls' clubs, life skills programmes and girls' well-being outcomes* (359–369). London: Overseas Development Institute.

Markus, H. R., & Kitayama, S. (2003). Models of agency: Sociocultural diversity in the construction of action. In V. Murphy-Berman & J. J. Berman (Eds.), *Cross-Cultural differences in perspectives on the self* (pp. 18–74). University of Nebraska Press.

Maxwell, J. A. (2004). Using Qualitative Methods for Causal Explanation. *Field Methods*, *16*(3), 243–264. https://doi.org/10.1177/1525822X04266831

Morris, Monique. (2019) *Sing a Rhythm, Dance a Blues.* The New Press.

Murphy-Graham, E., & Leal, G. (2015). Child Marriage, Agency, and Schooling in Rural Honduras. *Comparative Education Review*, *59*(1), 24–49. https://doi.org/10.1086/679013

Murphy-Graham, E., & Lloyd, C. (2016). Empowering adolescent girls in developing countries: The potential role of education. *Policy Futures in Education*, *14*(5), 556–577.

Oulo, B., Sidle, A. A., & Butler, M. (2021). *Stemming Pandemic-Related Losses in Girls' Education: Promising Practice from the AMPLIFY Collective* [AMPLIFY Girls Research Brief]. AMPLIFY Girls.

Oulo, B., Sidle, A. A., Kintzi, K., Mwangi, M., & Akello, I. (2021). *Understanding the Barriers to Girls' School Return: Girls' Voices from the Frontline of the COVID-19 Pandemic in East Africa* (AMPLIFY Girls Research Brief). AMPLIFY Girls.

Poteat, V. P., Calzo, J. P., & Yoshikawa, H. (2016). Promoting youth agency through dimensions of gay–straight alliance involvement and conditions that maximize associations. *Journal of Youth and Adolescence*, *45*(7), 1438–1451.

Psaki, S., Haberland, N., Kozak, M., & Woyczynski, L. (2021). *Girls' Education Roadmap: 2021 Report.* (EGER Reports). The Population Council.

Sahni, Urvashi. (2017). *Reaching for the Sky: Empowering Girls Through Education.* Brookings Institution Press: Washington, DC.

Scriven, M. (2008). A summative evaluation of RCT methodology: An alternative approach to causal research. *Journal of Multidisciplinary Evaluation*, *5*(9), 11–24.

Shah, P. P. (2016). Agency as negotiation: Social norms, girls' schooling and marriage in Gujarat, India. In J. DeJaeghere, J. Josić, & K. S. McCleary (Eds.), *Education and Youth Agency: Qualitative Case Studies in Global Contexts* (pp. 85–102). Springer.

Sidle, A. A., Ahigika, D., Aszed, W., Bashuna, S., Belle, S., Bideri, D., Bohart, E., Butler, M., Fortunate, K. H., Fowler, Z., Kamau, L., Littman, J., Makumayenzi, V., Mariki, E., Mirisho, A., Musumba, P., Omare, M., Omondi, D., Oulo, B., … Wolf, K. (2019). *Measuring Girls' Agency in East Africa—Co-Creating Contextually Specific Tools for Evaluation* (AMPLIFY White Paper). AMPLIFY Girls.

Sidle, A. A., & Oulo, B. (Unpublished Manuscript). Assessment of a Practitioner-Derived Framework for Measuring Girls' Agency in East Africa. *Comparative Education Review*.

Sidle, A. A., Belle, S., Bohart, E., Fortunate, K. H., & Fowler, Z. (Unpublished Manuscript). *Community-Driven Evaluation in Practice: Mapping Outcomes for Adolescent Girls in East Africa.*

Sperling, G. B., & Winthrop, R. (2016). *What works in girls' education: Evidence for the world's best investment*. Brookings Institution Press.

Sridharan, S., Ravindranath, P., Pottinger, E., & Cosentino, C. (2019). *Building Youth Life Skills: Lessons Learned on How to Design, Implement, Assess and Scale Successful Programming*. The Partnership to Strengthen Innovation and Practice in Secondary Education (PSIPSE) & Mathematica. https://drive.google.com/file/d/1-kV-XHhoxMUz1jyTTFgaRPK7eaE9ouO8/view.

Stuart, E. A., Bradshaw, C. P., & Leaf, P. J. (2015). Assessing the generalizability of randomized trial results to target populations. *Prevention Science : The Official Journal of the Society for Prevention Research*, *16*(3), 475–485. PubMed. https://doi.org/10.1007/s11121-014-0513-z.

Temin, M., & Heck, C. J. (2020). Close to Home: Evidence on the Impact of Community-Based Girl Groups. *Global Health: Science and Practice*, *8*(2), 300. https://doi.org/10.9745/GHSP-D-20-00015

Trochim, W. M., & McLinden, D. (2016). Introduction to a special issue on concept mapping. *Evaluation and Program Planning*. http://dx.doi.org/10/1016/j.evalprogplan/2016.10.006.

Trochim, W. M., Donnelly, J. P., & Arora, K. (2016). *Research Methods: The Essential Knowledge Base* (Instructor's Edition). Cengage Learning.

Trochim, W. M., & Linton, R. (1986). Conceptualization for planning and evaluation. *Evaluation and Program Planning*, *9*(4), 289–308.

Trochim, W. M. (1985). Pattern Matching,Validity, and Conceptualization in Program Evaluation. *Evaluation Review*, *9*(5), 575–604.

Unterhalter, E., Arnot, M., Lloyd, C., Moletsane, L., Murphy-Graham, E., North, A., Parkes, J., & Saito, M. (2014). *Interventions to Enhance Girls' Education and Gender Equality: Education Rigorous Literature Review*. Department for International Development London.

Vaughn, L. M., Jacquez, F., & McLinden, D. (2013). The use of concept mapping to identify community-driven intervention strategies for physical and mental health. *Health Promotion Practice*, *14*(5), 675–685.

Wallerstein, N., Duran, B., Oetzel, J., & Minkler, M. (Eds.). (2018). *Community-based participatory research for health: Advancing social and health equity* (Third edition.). Jossey-Bass, a Wiley Brand.

Wells, R., Wray-Lake, L., Plummer, J. A., & Abrams, L. S. (2021). *'I feel safe out here': Youth perspectives on community-based organizations in urban neighborhoods*.

CONCLUSION: "EM-PLACING" AGENCY

"*Emplace* verb (used with object), em·placed, em·plac·ing. **to put in place or position**: A statue was emplaced in the square."

Throughout my dissertation project I have repeatedly made the argument that agency is context-dependent. Perhaps a more useful way to understand the argument about the contextualized nature of agency is to consider it as 'emplaced,' or situated in a particular environmental context. I began my project with a clear hypothesis that agency was 'contextually specific'--that the specific abilities and attributes needed to create agentic capacity in girls was contextually-bound up in the opportunity structures and gendered norms of their communities and countries.

Over the course of my project, I have encountered other types of theorizing about *"*emplaced" agency, from authors who view agency as either relational,[29] [30] coalitional,[31] temporally situational,[32] disjoint/conjoint,[33] and/or distributed.[34] Ultimately, each of these conceptualization of agency (relational, coalitional, situational, distributed, etc.) also imply that agency is emplaced. For women and girls, agency is intimately tied to two aspects of context: socio-economic context—representing the types of material opportunities that exist for girls

[29] Dierker, B. (2016). "You Are Building on Something": Exploring Agency and Belonging Among African American Young Adults. In *Education and Youth Agency* (pp. 27–46). Springer.

[30] Kennelly, J. J. (2009). Youth cultures, activism and agency: Revisiting feminist debates. *Gender and Education*, *21*(3), 259–272.

[31] Chávez, K. R., & Griffin, C. L. (2009). Power, feminisms, and coalitional agency: Inviting and enacting difficult dialogues. *Women's Studies in Communication*, *32*(1), 1–11.

[32] Bajaj, M. (2009). 'I have big things planned for my future': The limits and possibilities of transformative agency in Zambian schools. *Compare*, *39*(4), 551–568.

[33] Markus, H. R., & Kitayama, S. (2003). Models of agency: Sociocultural diversity in the construction of action. In V. Murphy-Berman & J. J. Berman (Eds.), *Cross-Cultural differences in perspectives on the self* (pp. 18–74). University of Nebraska Press.

[34] Campbell, C., & Mannell, J. (2016). Conceptualising the agency of highly marginalised women: Intimate partner violence in extreme settings. *Global Public Health*, *11*(1–2), 1–16.

around school, employment and income; and social context—including and most especially, gender norms, gendered expectations and histories of patriarchy. [35]

Understanding these constraining aspects of context, and specifically the relationship of agency to gendered realities, re-affirms the importance of thinking about agency and its practice in partnership with those who best understand social settings where girls' live and grow. I urge more work that centers community voices—both practitioners on the front line and girls themselves—and that digs deeper into the questions about program practice as an 'emplaced' undertaking. Such approaches are needed in order to avoid the common pitfalls of replicating programs but not outcomes, or measuring soft skills based on de-contextualized definitions which hold little validity in local settings. Achieving quality education goals for girls around the world requires that our approaches to measurement and practice must be attentive to girls' unique constraints, and opportunities, and the variation of these across place.

In paper one, I have argued how agency even in its original conceptualization in the human development literature, is necessarily tied to an individual's environment. I offered in paper two, a practical example how to include environmental context in the measurement of agency through participatory processes of survey development. In paper three, I argue why 'place' doesn't just matter for theorizing about the psychological dimensions of a construct, but why the practice of agency necessitates context-inclusive theories which are more likely to replicate outcomes than traditional attempts at scale. "Emplacing" agency is both an approach to a particular problem of practice in girls' education, but also a wider call to action. Global approaches to youth success must reconsider the old ways of thinking about current problems of

[35] Willan, S., Gibbs, A., Shai, N., Ntini, N., Petersen, I., & Jewkes, R. (2020). Did young women in South African informal settlements display increased agency after participating in the Stepping Stones and Creating Futures intervention? A qualitative evaluation. *Social Science & Medicine*, *265*, 113302.

practice in education as challenges of ‘scale’ and ‘consensus’ but as problems of contextualization and conceptualization for greater impact.