Evaluation Education in a World in Transformation: The Way Forward

Marcia Paterno Joppert
Claremont Graduate University

Follow this and additional works at: https://scholarship.claremont.edu/cgu_etd

Part of the Educational Assessment, Evaluation, and Research Commons, and the Psychology Commons

Recommended Citation

This Open Access Dissertation is brought to you for free and open access by the CGU Student Scholarship at Scholarship @ Claremont. It has been accepted for inclusion in CGU Theses & Dissertations by an authorized administrator of Scholarship @ Claremont. For more information, please contact scholarship@claremont.edu.
Evaluation Education in a World in Transformation: The Way Forward

By

Marcia Joppert

Claremont Graduate University

2023
Approval of the Dissertation Committee

This dissertation has been duly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of Marcia Joppert as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy.

Dr. Stewart I. Donaldson, Chair
Claremont Graduate University
Distinguished University Professor

Dr. Brad Cousins
University of Ottawa
Emeritus Professor

Dr. Tiffany Berry,
School of Social Science, Policy & Evaluation
Dean
Claremont Graduate University
Full Research Professor

Dr. John LaVelle
University of Minnesota-Twin Cities
Assistant Professor
The world has experienced rapid changes, leading to pressing issues such as environmental degradation, social inequality, and resource depletion. As a transdisciplinary field, evaluation has emerged as a crucial tool in addressing these challenges and promoting systemic change. However, concerns have been raised regarding the field's capacity to meet these expectations and the existing gaps in evaluation education. This research aims to address these gaps by exploring how formal evaluation education programs (EEPs) respond to the preparation of professionals for systems change evaluations, prompting discussions on the need to redefine the approach to teaching evaluation. The study adopts a sequential multi-phased mixed-methods design, comprising two qualitative and one mixed-methods study.

Study 1 involves semi-structured interviews with 19 international evaluation experts from different roles, locations, and backgrounds to gather their perspectives on the role of evaluation in supporting systems change. Critical topics, approaches, and methods that should be included in formal EEPs to prepare professionals for this challenge are also pointed out. The findings highlight that evaluation can contribute to systems change by providing sound evidence and recommendations for problem understanding, solution design, progress tracking, and strategy adjustments. However, influential contributions require adopting interdisciplinary and systemic
approaches that respond to diverse contexts, voices, and considerations of equity and sustainability. Moreover, evaluations must stay abreast of new technologies for data collection, collaboration, and continuous learning. The insights from Study 1 identified desirable traits for EEPs, encompassing curricula design, modalities, strategies for developing soft skills, and connections to the broader professional field of evaluation.

Study 2 involves a focus group with 11 young and emerging evaluators (YEEs) from different parts of the globe to explore their perspectives on the role of evaluation in supporting systems change and their challenges, needs, and interests regarding evaluation education. The study reveals that YEEs recognize the importance of evaluation in driving systems change, emphasizing the need to shift from traditional linear models to systemic thinking. YEEs advocate for accessible and inclusive EEPs that provide real-world practical opportunities and instruction in various evaluation theories. They recommend incorporating utilization-focused evaluation, people-centered and culturally responsive approaches, systemic evaluation methods, and foresight techniques to anticipate future transformation scenarios. Including new data collection methods like storytelling and artificial intelligence, as well as incorporating youth perspectives in evaluation processes, is also suggested. Studies 1 and 2 contribute to identifying the anticipated characteristics that will shape the future trajectory of evaluation.

Study 3 compiles a comprehensive list of formal postgraduate EEPs offered by academic institutions worldwide. Through descriptive and categorical statistical analysis, these programs are assessed to determine the extent to which they incorporate the characteristics identified in Studies 1 and 2. The analysis reveals significant shortcomings in integrating the relevant topics,
approaches, and methods suggested by the previous studies and a weak link with professionalization. A total of 271 formal postgraduate programs from 39 countries across all continents and offered in 9 languages were identified and analyzed, including 45 doctorates, 124 master's, 58 certificates, and 44 diplomas. The data collected in Study 3 were shared with the Global Evaluation Initiative (GEI) to create and publicly disseminate a comprehensive global directory of postgraduate academic programs through the "Better Evaluation" knowledge platform.

This research underscores the importance of critically examining the alignment between evaluation trends and EEPs on a global scale. Furthermore, it provides novel insights by offering specific recommendations for integrating emerging content into evaluation education to better prepare students for the demands of the field. The findings offer valuable guidance for future students in selecting EEPs, contribute new elements to research on evaluation education, and encourage EEP leaders to reflect on the relevance and effectiveness of their offerings. Ultimately, this research emphasizes the significance of comprehensive evaluation education.
Dedication

To my late mother, Maria Jose Paterno Joppert, who remains a great reference for me in terms of honesty, courage, and sensitivity.

To my husband, Laerte Rimoli, who has always been ready to embark on life's adventures with me and has provided unwavering support, patience, and dedication over the past six years.

To my son, Fernando Rimoli, whom I hope will continue to follow the path of knowledge, curiosity, science, and ethics.

To Ana Elisa Tavares de Miranda, the godmother of the entire family.
Acknowledgements

I could not complete such an important cycle in my life without expressing my profound gratitude to the people who have accompanied me since the very first idea of pursuing a PhD at the age of 53. Obtaining a PhD degree is a great achievement, given that only 0.20% of people in my country, Brazil, have had this opportunity. Mostly, I am the first person in my entire family to achieve this feat. I wanted to say that this is not about getting a degree to make more money. Instead, my greatest motivation is my passion for this still in development field. To the readers of this dissertation, know that I am very emotional to have come this far and that the most important thing for me is the potential contribution that my work will give to the evaluation field. My profound gratitude goes to:

The Coordination for the Improvement of Higher Education Personnel of the Ministry of Education (CAPES Foundation) for believing in my ability to deliver something important for my country and for granting me a scholarship. All the fellows who have accompanied me in the leadership of the Brazilian Monitoring and Evaluation Network for the last 15 years, for the opportunity to reflect and build relevant actions for the field in Brazil. CGU for accepting and supporting me throughout the Ph.D. journey, which made me grow incredibly. Special thanks to my advisors, Stewart Donaldson, Brad Cousins and Tiffany Berry, and to my professors Leslie Fierro, Tarek Azzam, Tiffany Berry, Michael Patton, Andrew Conway, William Crano, and Kathy Pezdek. Finally, gratitude for John LaVelle for accepting to be the external examiner of my dissertation defense.

To Jos Vaessen, Varsha Thebo and the Global Evaluation Initiative IT team, for making my data useful for the broad evaluation community.
To Mr. Jim Kelly, my beloved English teacher, who not only improved my language skills but became my best friend.

To Caique Novis, who showed me that life always has more than one possibility.

To Humberto Martins, partner, and encourager for a whole lifetime.

To Emeritus Professor Evert Vedung, for the gentle recommendation.

To Thereza Penna Firme, one of my inspirations.

To my Southern Californian "family": Marina Zanzarini, Pedro Diniz, Anabella Diniz, Lucca Diniz, Fernanda Konomi, Jennifer Salas Feitosa, Diego Salas, Luca Feitosa, Barbara Blatz Stone, Larissa Loures, Luigi Markzac, Luan Markzac, Leandro Markzac, Dana Lopez, William Granja, Tiffany Granja, Michael and Bruna Fileta, Angela and Steve Scheibe, Cristiane and Ed Roedel, Sara Hirata and Gabriel Granco, and all the others, not mentioned here, who always had a smile and a word of affection that were fundamental to encourage my journey and maintain mental health especially during pandemic times.
Table of Contents

Chapter 1: Introduction

Chapter 2: Literature Review
  Evaluation as a Profession
  A World in Transformation and the Implications for Evaluation Practice and for Evaluators
  The Way Forward
  The Current Study
  Research Questions

Chapter 3: Methods
  Research Design and Overview
  Phase 1
  Phase 2

Chapter 4: Results
  Study 1: Interviews with International Evaluation Experts
  Study 2: Focus Group with Young and Emerging Evaluators (YEEs)
  Study 3: Evaluation Education Programs

Chapter 5: Conclusion

Chapter 6: Discussion
  Strengths
  Limitations
  Implications and Future Directions

References

Appendix A
Appendix B
Appendix C
Appendix D
Appendix E
Appendix F
Appendix G
Appendix H
Appendix I
List of Figures and Tables

Figure 1: *Overview of Methods and Design*

Table 1: *Initial List of Codes for Interviews Analysis*

Table 2: *Initial List of Codes for Focus Group Analysis*

Table 3: *Summary of Evaluation Education Programs Across the World (N=271)*

Table 4: *Aspects of EEP Analysis and Rating System*

Table 5: *Coding System Adopted to Designate the Interviewees Based on their Role within the Evaluation Domain and Geographical Region of Origin*

Table 6: *Core Curriculum Dosage Level Across Types of Programs (N=271)*

Table 7: *Core Curriculum Dosage Across Regions (N=271)*

Table 8: *Core Curriculum Dosage Level Across Disciplines (N=271)*

Table 9: *Core Curriculum Dosage Level Across Modalities (N=271)*

Table 10: *Programs Offering Elective Classes Across Types of Programs (N = 271)*

Table 11: *Programs that Declared or Suggested the Inclusion of Topics Identified as Potentially Crucial for Evaluation Practice in the Future (N = 271)*

Table 12: *Topic Dosage Level Across Disciplines (N = 271)*

Table 13: *Number of programs by Approach Level (N = 271)*

Table 14: *Relevant Approaches Declared or Suggested*

Table 15: *Approaches Dosage Across Topics Dosage (N = 271)*

Table 16: *Programs that Declared or Suggested the Inclusion of Specific Research Methods in the Curricula*
Table 17: Inclusion of New Technologies in the Curricula Across Regions (N = 271)

Table 18: Practicum Offered Across Regions (N = 271)

Table 19: Inclusion of Professionalization Across Disciplines (N = 271)

Table 20: Summary of Chi-square Tests of Independence Results
Chapter 1: Introduction

Evaluation has had many published definitions over the past thirty years. One of the first, still the most cited, was made by Michael Scriven:

Evaluation refers to the process of determining the merit, worth, or value of something, or the product of that process. Terms used to refer to this process or part of it include: appraise, analyze, assess, critique, examine, grade, inspect, judge, rate, rank review, study, test…The evaluation process normally involves some identification of relevant standards of merit, worth, or value; some investigation of the performance of evaluands on these standards; and some integration or synthesis of the results to achieve an overall evaluation or set of associated evaluations. (Scriven, 1991, p. 139).

Although evaluation is a broad term that can be applied to various objects or "evaluands," it typically involves a systematic process of identifying criteria or principles for assessing value or worth, comparing the evaluand to those criteria, and arriving at conclusions that inform decision-making. While this definition is commonly accepted in academia, it is not the only one found in the literature. So, there is no consensus about the use of a standard definition or interpretation of evaluation in the international development sector, and the terms "evaluation" and "monitoring" are often used interchangeably, leading to confusion and a shift in focus from "learning from the program" to "tracking of the program." This has resulted in evaluation activities being narrowly employed to assess whether all program activities are completed as planned rather than understanding the merit and value of the program for its stakeholders and broader society. (Dighe & Sarode, 2019)

In the United States, the concept of evaluation as an organized practice has sparked discussions about whether it qualifies as a knowledge occupation or a profession. Despite these
debates, there is no consensus among scholars regarding whether evaluation should be classified as a profession (Patton, 1990; Piciotto, 2011; Schwandt, 2015). Nevertheless, the evolution of evaluation practice on a global scale demonstrates a clear progression toward professionalization.

The education of preservice evaluators is crucial for the maintenance and growth of program evaluation as a profession (Gullickson et al., 2019). A recent study conducted by King and Ayoo (2020) surveyed evaluation education programs across four regions, namely the U.S., Europe, the Asia-Pacific region, and Canada from 1978 to 2018. The analysis of the curriculum dimension revealed various gaps, including an imbalance between technical skills and soft skills such as project management, ethics, communication, and conflict resolution, a lack of supervised internships, and a discrepancy between the skills that employers sought and what novice evaluators learned in their courses. Additionally, Gullickson et al. (2019) highlighted challenges associated with transdisciplinarity. While some scholars support the idea of evaluation as a transdisciplinary field (Donaldson & Scriven, 2003; Patton, 2019), it is challenging to find education programs outside thematic disciplines (LaVelle, 2014; Gullickson et al., 2019). King and Ayoo (2020) found that the lessons learned from their study echoed a statement made 40 years ago: “There is no answer to the question [of] what constitutes an adequate evaluation training program” (Brown, 1980, p. 86, cited by King & Ayoo, 2020).

The evaluation field acknowledges the significance of adequately training evaluators and has made considerable progress in this area. Nevertheless, there remains a certain level of doubt regarding the proficiency of those who conduct evaluations (Rodríguez-Bilella, 2017). Although some countries, such as Canada, Japan, and Taiwan have introduced credentialing procedures, there is limited evidence to support the notion that these processes have substantially enhanced the quality of evaluators’ work. Over the past decade, the rise in expectations and demands for
evaluations of intricate subjects, such as global agreements, has engendered the notion that evaluations can contribute to addressing critical global issues.

To address global problems affecting the planet, mobilization of society as a whole is required. The international development community has established and advocated for global agreements to tackle critical issues such as poverty, inequality, social injustices, and sustainability. The Agenda for Sustainable Development, also known as Agenda 2030, was launched in 2015 by the United Nations (UN) and signed by 195 UN member countries. It comprises 17 Sustainable Development Goals (SDGs) that balance the three dimensions of sustainable development: economic, social, and environmental. Agenda 2030 also includes 169 targets and 203 indicators at different levels, from global to local. Although ambitious, Agenda 2030 offers a systemic view of transforming the world we have into the world we want and raises questions about societal choices (Desa U.N., 2016). Monitoring and evaluation are strategically important to track and communicate progress and evaluate impacts, as Agenda 2030 has a transparent follow-up mechanism. In 2015, the UN declared it the International Year of Evaluation (EvalYear) following a strong advocacy campaign. Since then, there has been a growing discussion about the role of evaluation in addressing global issues affecting the planet. This new role will require changes in how evaluation is practiced and in the competencies that evaluators (both experienced and novice) will need to acquire to respond to new types of expectations.

The purpose of this research is to investigate how formal evaluation education programs can adapt to the evolving challenges in the field of evaluation in a rapidly changing world. To achieve this, the study gathered perspectives from international experts in evaluation and from young and emerging evaluators on the themes, approaches, and methods that should be
incorporated into evaluation education programs. Finally, the study compared the content of a comprehensive sample of current evaluation education programs with the key themes, approaches, and methods identified in the previous phases, highlighting similarities and differences. After the introductory chapter, Chapter 2 provides a comprehensive literature review of the professionalization of evaluation, contextualizing the changes that the world has undergone over the past two centuries and examining the implications and challenges these changes pose for evaluators and evaluation practice. The chapter concludes with a discussion of the competencies that evaluators must acquire to meet the new and future challenges of the field. Chapter 3 outlines the research methods employed, including the participants, instruments, data collection procedures, and analysis plan for the three studies conducted. Chapter 4 presents the results of each study and their key takeaways. Chapter 5 presents the research conclusions, and finally, Chapter 6 discusses the strengths and limitations of this research.
Chapter 2: Literature Review

The primary goal of the literature review was to examine the landscape of knowledge base and identify gaps related to aspects of analysis of evaluation education. Specifically, the review aimed to identify the perspectives of experts on the applicability of professionalization concepts to evaluation and the primary factors that can impact it. For each factor of influence, recent literature was selected and analyzed, with a focus on empirical studies. The second section of the review delves into the historical context of transformative events that have impacted the world since the Industrial Revolution and their implications for the evaluation practice and evaluators. Finally, the review considers the path ahead, specifically the competencies that evaluators must develop or enhance to ensure that evaluations can contribute effectively to system change.

Evaluation as a Profession

There is ongoing debate as to whether evaluation should be considered a profession. Since the emergence of evaluation as an organized practice in the U.S., the question of whether it is an occupation or a profession has been a subject of discussion. Picciotto (2011) identified six key characteristics of professions out of the 23 identified by Millerson (1964) among 21 authors: possession of a skill based on theoretical knowledge, provision of training and education, testing of members’ competence, professional organization, adherence to a code of conduct, and altruistic service. This study reviewed recent literature to analyze how closely the practice of evaluation aligns with these six characteristics.

Possession of a Skill Based on Theoretical Knowledge

The field of evaluation is rich in literature that includes theoretical frameworks, empirical research, and practical applications. A vast array of peer-reviewed journals dedicated to
The Voluntary Organizations for Professional Evaluation (VOPEs), comprising evaluation associations, societies, and networks worldwide, are critical in disseminating knowledge and practices through organizing conferences, forums, and dialogues. These events serve as opportunities for evaluation professionals to exchange ideas, share experiences, discuss emerging trends, and consider the field’s future challenges. Knowledge sharing is also facilitated by online mechanisms such as social media groups and electronic lists, which are gaining popularity (Rodríguez-Bilella & Lucero, 2016).

**Provision of Education**

The continued development and expansion of evaluation practice rely on effective education programs that prepare skilled and motivated professionals capable of producing high-quality evaluations across all disciplines and meeting clients’ expectations (Gullickson et al., 2019). Access to education and training can be attained through various avenues and may be influenced by individual factors such as geographic proximity, financial affordability, and the ability to balance academic pursuits with personal and professional obligations (LaVelle, 2014).

Articles about evaluation education were first published in the late 1970s, marking the beginning of a new area of study on the preparation of evaluators (King & Ayoo, 2020). Since 1976, research on evaluation education has been ongoing, with the first evaluation education directory compiled by Gephart and Potter, identifying 80 evaluation education programs. These programs ranged from single courses to structured curricula aimed at preparing "methodologically strong evaluation practitioners" (LaVelle & Donaldson, 2015, p.43). In 2014, LaVelle identified 43 certificate programs and 81 university-based evaluation education programs (MA, MS, Ed.S, and Ph.D.) in the United States and overseas. These programs and
degrees were mainly housed in education, educational psychology, psychology, and public policy departments, and were English-only. Other regions have also been mapping education opportunities in evaluation. For example, a recent study identified more than 20 graduate programs in evaluation in Latin America and the Caribbean (Ghiano et al., 2017), while another study discussed European university-based programs in evaluation and their future challenges (Friedrich, 2016). African evaluation graduation programs were also mapped recently (Basheka & Byamugisha, 2015; Wildschut & Silubonde, 2020). Additionally, Hung et al. (2012) explored the lack of training in the Asia-Pacific region.

Three decades ago, Patton (1990) emphasized that for evaluation to thrive as a profession, it required vision, quality products and processes, and skilled, trained evaluators. At the time, Patton argued that training needed to move beyond multiple and diverse methods and techniques to focus on relevant skills such as conceptualization and program logic, interpersonal relationships, political sensibility, creativity, and communication. A decade later, Stufflebeam (2001) also stressed the importance of evaluation education and training programs to provide a continuous flow of excellently qualified and motivated evaluators for the evaluation field’s future success. Recent research further underscores the critical importance of preservice education for evaluators to maintain and grow program evaluation as a profession (Gullickson et al., 2019).

The evaluation field is diverse, and so are the professionals working in it. For some, evaluation is a specialty pursued in academic settings with opportunities for research or teaching, while for others, it is a professional career in applied science with employment opportunities in public, for-profit, or non-profit organizations. Gullickson et al. (2019) applied the Context, Input, Process, Product (CIPP) model (Stufflebeam & Zhan, 2017) to analyze evaluation education
programs. Their analysis revealed potential learner groups with possible overlaps, including evaluation practitioners, individuals in academic settings, accidental evaluators, program consumers or beneficiaries, under-represented groups, and commissioners who are responsible for contracting and supervising evaluation services. They concluded that “the intended outcomes of evaluator education are to enable learners to develop (a) awareness of evaluation as a profession; (b) evaluation knowledge, skills, and dispositional attributes, including the ability to adhere to ethical practice; and (c) the ability to recognize and deliver quality evaluation practice. “(Australasian Evaluation Society, 2013, cited by Gullickson et al., 2019, p. 26).

Accessing education and training in the field of evaluation can be done through various avenues, depending on factors such as location, cost, and personal and professional responsibilities. One pathway is formal education through certificate, master’s, and doctoral degree programs. Another option is through training modalities such as professional development workshops, webinars, and on-site training opportunities offered by VOPEs, consulting companies, and university-affiliated centers (LaVelle, 2014). These options are particularly useful for "accidental evaluators," who need to acquire basic knowledge, or for those interested in learning about evaluation due to its connection with their current roles (Schwandt, 2015). The trend of developing evaluators’ skills through rapid modalities aligns with the goal of disseminating basic knowledge among professionals in various fields, including health, education, and social services to meet the demand for evidence. However, this trend can also lead to evaluators being viewed solely as technical experts equipped with tools for results-based management, program monitoring, performance assessment, and impact evaluation, which provide reliable evidence of outcomes. While providing evidence is essential, there is a risk of limiting the evaluation function to merely verifying and reporting the achievement of targets and
outcomes, which undermines the primary purpose of evaluation — to contribute to the public good, the original calling of the field (Schwandt, 2015).

Evaluation education programs suffer from several gaps. King and Ayoo (2020) conducted a comprehensive literature review of 64 publications from 1978 to 2018 on evaluator education. The review highlighted the expansion of university-based and certificate programs. The publications were categorized into research studies and reflective narratives, which covered four parts of the world, including the U.S., Canada, Europe, and the Asia-Pacific region. The review analyzed evaluator education based on four dimensions: evaluation students, evaluation faculty, evaluation education programs, and evaluation education curriculum. The curriculum dimension analysis showed gaps in the programs. Firstly, the systematic inquiry domain was present in most evaluator education programs, followed by professional practice. Technical skills like data collection and statistics were more prevalent than project management, ethics, communication, conflict resolution, and the management of multiple projects. Students reported learning the latter set of competencies through professional development workshops. Secondly, the review identified a lack of supervised internships, which are crucial for learners to gain practical experience and soft skills. Thirdly, novice evaluators who graduated from master’s and doctoral programs reported that they were not taught or did not learn the skills employers were looking for, including interpersonal skills, report writing, project and team management, practical research design, and evaluation theory. These findings echo a statement from 40 years ago that "there is no answer to the question of what constitutes an adequate evaluation training program" (Brown, 1980, p. 86, cited by Gullickson et al., 2019).

Numerous authors have highlighted gaps in evaluator education. Meyer (2016) contends that there is a conflict between academic education and applied evaluation practice, as academic
teaching adheres to academic values and standards, whereas evaluation practice must consider market laws, time constraints, and customer demands. Gullickson et al. (2019) identify challenges in evaluation education, including the need for multidisciplinary training, which is scarce outside thematic disciplines. LaVelle’s (2014) study revealed that most university-based evaluation education programs are housed in education, educational psychology, psychology, and public policy departments, which affects their frameworks, methods, and theories. Another challenge is the absence of a common set of agreed-upon competencies for program evaluators. While some countries, such as Canada, have developed competencies and skills for evaluators linked to education programs, this is not common in other countries. Furthermore, there is no specific accreditation process for evaluation education programs. Schwandt (2015) suggests that professional schools should have accreditation processes that connect standards, competencies, and curriculum, but this does not exist in evaluation.

Ensuring that the current diverse range of educational opportunities is sufficient to equip professionals with the skills needed to conduct and use rigorous, visible, and responsive evaluations, as well as teach and research in the field, appears to be a daunting task. Considering this, Gullickson et al. (2019) and King and Ayoo (2020) recommend that education efforts adopt a continuous evaluative lens to enhance teaching and evaluation practices. Ultimately, more work is required to advance evaluator education to better meet the learning and performance demands of the global context, as Gullickson et al. (2019) conclude.

**Testing of Members’ Competencies**

The recognition of evaluators’ competencies is currently a significant issue in the evaluation community. Although the evaluation field acknowledges and values the importance of preparing evaluators, skepticism remains about the quality of professionals who conduct
evaluations (Rodríguez-Bilella, 2017). Accurate assessment of evaluators’ knowledge, skills, and experience is necessary, particularly in the selection processes of faculty, scholars, independent consultants, or external evaluation teams. Different types of criteria have been used to assess evaluators’ competencies, including analyzing educational backgrounds, portfolios, technical proposals, and interviews. Some professional associations have defined context-specific sets of competencies to support professionalization in their countries. In addition, formal recognition mechanisms for credential evaluators have been established by evaluation associations in Canada, Japan, and Thailand (King & Stevahn, 2015). However, research shows that the high costs and time associated with the credentialing process represent a barrier for evaluators and there is no substantial degree of acceptance and support from commissioners and contractors (Fierro et al., 2016). Even in places where models for credentialing evaluators have been tested, such as in Canada, the path to professionalization is not yet evident. Therefore, it is essential to move towards recognizing the value of evaluation and the importance of conducting it professionally.

Evaluators must take steps to be recognized by the public or the market as possessing special knowledge and skills and being prepared to apply that knowledge and skills in the interests of others (Rodríguez-Bilella, 2017). Despite the potential benefits, some members of the American evaluation community have expressed concerns about adverse side effects, such as excluding already experienced evaluators based on their background or lack of certain competencies (Donaldson, 2019). Thus, establishing a process for credentialing evaluators requires careful consideration of its benefits and limitations.

**Professional Organization**
In the last thirty years, the importance of the professionalization of evaluation has been recognized by evaluators. The demand for program evaluation in the U.S. began in the 1960s when President Kennedy’s administration launched comprehensive social programs. By 1979, approximately $428 billion had been invested in these programs in the U.S., which raised concerns in Congress about accountability, side effects, mismanagement, and the legitimacy of social welfare. To demonstrate that funds had been spent properly and with desirable results, advocates of social programs required program evaluation. However, the government lacked personnel with the necessary skills to fulfill this demand. Initially, professional accountants, management consultants, and research and development specialists were hired to meet this need. However, by the late 1960s, demand for evaluation surpassed the supply of skilled professionals (Shadish et al., 1991). This gap in supply led to the creation of the Canadian Evaluation Society (CES) in 1981, followed by the American Evaluation Association in 1986.

These two organizations served as models for the establishment of national and transnational organizations. VOPEs have grown significantly, from 15 organizations in the 1990s to more than 155 by early 2013, with an estimated membership of 34,000 (Rugh & Segone, 2013). According to the International Organization for Cooperation in Evaluation (IOCE) VOPEs Directory, there are currently 145 VOPEs with approximately 40,000 members worldwide, including national, regional, and international organizations. By joining VOPEs, evaluation professionals commit to their values and uphold their ethical principles, standards, and competencies. Additionally, they have access to resources, information, networking opportunities, education, and job prospects.
Adherence to a Code of Conduct

The provision of guidelines and standards for evaluation is a crucial ethical foundation that directs professional practice (Bustelo, 2006). Ethics involves making decisions based on one’s beliefs about what is right. Evaluation practitioners frequently encounter ethical dilemmas, and in response, many VOPEs have established ethical principles and guidelines. Rather than offering prescriptions for resolving ethical dilemmas, these guidelines are intended to inform evaluators of components of professional integrity (Schwandt, 2015). For instance, a study by Morris and Clark (2013), involving a sample of 925 non-student members of the AEA explored evaluators’ experiences with pressure from clients to misrepresent evaluation findings to make them look more favorable than in the original report. The study found that 42% of respondents reported encountering this type of situation, and 70% of these respondents indicated that it had occurred on more than one occasion.

Altruistic Service

The 145 VOPEs worldwide are defined as "non-profit membership organizations open to individuals interested in evaluation" (IOCE, [n.d.]). As voluntary organizations, VOPEs do not benefit any individual or corporate entity and aim to leverage the culture, disseminate knowledge and information, and promote the professionalization of the evaluation field in their area of influence in an altruistic manner. These characteristics may vary among regions and countries and depend on each context. This diversity of contexts may be one reason why scholars have disagreed on whether evaluation constitutes a profession or not (Patton, 1990; Piciotto, 2011; Schwandt, 2015). Nonetheless, the global trajectory of evaluation is clearly advancing towards professionalization. Having described and considered issues and developments associated with
the professionalization of evaluation, the literature review now turns to the context within which evaluation has operated since it started to be practiced as a profession.

A World in Transformation and the Implications for Evaluation Practice and for Evaluators

A World in Transformation

Since the onset of the Industrial Revolution, the world has undergone notable transformations that have had both positive and negative consequences. Prior to this period, economies were rooted in social relations and governed by social norms and customs. Land, labor, and money were viewed as gifts of nature, human beings, and the state, respectively. However, the Industrial Revolution marked the start of the modern era, where the market began to play a central role in society’s values. The rise of capitalism brought about the commodification of land, labor, and money, and led to the transformation of social relations and economic structures (Polanyi, 2001). This process of commodification generated movements that aimed to protect societies from negative social consequences such as worker exploitation and inequality, as well as mitigate their effects. As a result, organizations such as trade unions, cooperatives, and social welfare programs emerged.

The Industrial Revolution marked the beginning of the Anthropocene era, where humans have become the dominant force shaping the planet. Technological innovations resulting from progress allowed humans to extract resources and exploit the environment on an unprecedented scale. The overuse of resources has led to a variety of environmental problems, including climate change, deforestation, and pollution (McKibben, 2006). Mass extinction of species is taking place on Earth as a result of human activities, particularly the burning of fossil fuels, deforestation, and overfishing. While previous extinction events were caused by natural events
such as volcanic eruptions or asteroid impacts, the current extinction event is being caused by human activities, which are causing rapid changes to the planet’s climate, habitats, and ecosystems (Kolbert, 2014). According to the author, society must take immediate action to address the root causes of this crisis and protect the many species that are currently under threat.

The Industrial Revolution ushered in progress without consideration for the long-term consequences of overusing human and natural resources, causing resource depletion, social inequality, and environmental degradation among other challenges, and there is a common understanding that this model is unsustainable (Wright, 2004). In addition to the Industrial Revolution, digital technology advances have also had a significant impact on society and the economy. Big tech companies such as Google, Meta, and Amazon can monetize data they collect from users in unprecedented ways, generating the so-called "surveillance capitalism." This transformation has significant implications for privacy, democracy, and human autonomy.

The United Nations Conference on Environment and Development in 1992, held in Rio de Janeiro, Brazil (Eco-92) led to the establishment of various international treaties, including Agenda 21, which defined eight Millennium Development Goals for sustainable development. After the end of Agenda 21’s deadline in 2021, the UN proposed a new Agenda for Sustainable Development called Agenda 2030, which includes 17 Sustainable Development Goals aimed at addressing global issues and mobilizing all segments of society to achieve those goals. Agenda 2030 represents a systemic view for transforming our world and raises questions about our societal choices. The SDGs encompass a range of areas, such as eliminating poverty and hunger, promoting health, education, gender equality, clean energy, sustainable cities, and communities, and fostering peace, justice, and strong institutions. Agenda 2030 challenges UN member countries to act and shift the responsibility for impact evaluations from donor organizations to
national, local, and non-governmental organizations. Monitoring and evaluation were recognized as functions of strategic importance to track progress and evaluate the impacts of the SDGs. The UN designated 2015 as the International Year of Evaluation after a robust advocacy campaign. (Catsambas & Bauer, 2015).

The concept of sustainable development at the local level necessitates policymakers “to correlate actions with specific changes in human and environmental conditions, test the strategic hypotheses (choice of priority issues, analysis of underlying factors, prescription of actions), assure accountability, capture lessons and develop capacity through learning” (Dalal-Clayton/Bass, 2002, cited by Meyer, 2020, p.222). Policy formulators must establish national monitoring and evaluation systems “to track systematically the key variables and processes over time and space and see how they change as a result of strategy activities” (p. 309). The Sustainable Development Goals (SDGs) encompass a follow-up and review mechanism that operates at multiple levels, which are voluntary and country owned. This mechanism should be inclusive, transparent, and participatory, while simultaneously being grounded in rigorous evidence (Meyer, 2020).

The evaluation community has been discussing the role of evaluation in addressing global issues for the past seven years, with a focus on integrating sustainability issues, supporting Agenda 2030, and promoting systemic and transformational change. The Prague Declaration was signed in 2019, acknowledging the urgent need for systemic change to address global crises. Recent conferences by the American Evaluation Association and European Evaluation Society have focused on staying relevant to critical issues, contributing to the improvement of others’ conditions, and transforming evaluation practice to meet complex challenges. The main goal of
the 2021 European Evaluation Society Conference was to identify concrete actions in four themes related to necessary paradigmatic shifts and transformation.

Large-scale initiatives such as the Feed Africa program have arisen to tackle complex global challenges and drive system transformation. The Feed Africa program is an initiative from the African Development Bank that aims to transform African agriculture into a competitive and inclusive agribusiness sector, with objectives that include ending extreme poverty, eliminating hunger and malnutrition, becoming a net exporter of agricultural commodities, and reaching the top of key agriculture value chains (African Development Bank Group, 2016). Evaluation is a crucial factor in addressing these challenges and contributing to the global system’s transformation. Two evaluations conducted on the Feed Africa initiative for the African Development Bank and the International Fund for Agricultural Development revealed that significant changes in organizational culture are necessary when working on value chains (Felloni & Kumbi, 2021). The last Annual Global Sustainable Development Report reveals that many countries are not on track to meet the 167 targets established for the 17 SDGs. The COVID-19 pandemic has exacerbated these challenges, impacting health, education, poverty, the economy, employment, gender issues, and peace. It highlights the complexity, inequality, and unpredictability of the world we live in.

The COVID-19 pandemic and Agenda 2030 highlight the importance of using evidence to make policies and decisions when dealing with complex and changing situations, where millions of lives are at stake. They have also raised awareness about the need for systemic evaluation that covers a broader range of topics beyond social areas such as education, health, and social assistance. The evaluation community has recognized the importance of applying evaluation concepts and methods to areas that have not been explored, such as climate change.
The year 2020 has shown the significance of using evidence to learn from the past and predict the impacts of future events. The "Prague Declaration on Evaluation for Transformational Change" has recognized the importance of solid support for transformational evaluation, which should be undertaken in partnership, promoting inclusiveness and respecting rights and responsibilities (Chaplowe and Hejnowicz, 2021). Having contextualized the dynamic global environment in which evaluation has emerged as a professional activity, the literature review now examines the ramifications of these contextual shifts on the practice of evaluation.

**Implications and Challenges for Evaluation Practice**

Meyer (2020) highlights three primary challenges in evaluating systems change initiatives like the Agenda 2030. The first comes from acknowledging the complexity and applying a holistic perspective to fully understand Agenda 2030, because the three dimensions — social, economic, and environmental — have many interlinkages. A second challenge embedded in Agenda 2030 is related to the idea of transformation to sustainable systems. Thirdly, the principle of inclusion, expressed through the slogan "leaving no one behind," poses a challenge to evaluating the effectiveness of the agenda’s implementation.

**Complexity and Systems Thinking.** Bamberger et al. (2016) presented a framework designed to assist with comprehending the complexities associated with evaluating interventions. The authors identified five fundamental sources of complexity that should be considered when assessing interventions: (1) the comprehension of the initiative, (2) the comprehension of the contextual environment in which the initiative operates, (3) the identification of the primary stakeholders involved and their interactions, (4) the nature of change and causality and how to capture it, and (5) the evaluation process itself. This approach is particularly useful for evaluating systems change initiatives because it considers multiple factors and stakeholders that interact in
dynamic and unpredictable ways. By taking a holistic approach and considering these different sources of complexity, evaluators can gain a more nuanced understanding of the intervention and its effects, which is essential for assessing systems change initiatives that aim to transform complex systems. The framework can also aid evaluators in designing evaluation processes that are sensitive to the unique challenges and complexities of systems change initiatives.

Despite the existence of numerous approaches to systems thinking, Gates et al. (2021) identified four distinct meanings associated with this concept. The first meaning involves focusing on the whole rather than the individual parts of a system, which provides a contrasting perspective to methods that reduce phenomena to separate parts to understand each component’s functioning. The second meaning involves the application of system theories, methods, approaches, or tools. The third meaning characterizes systems thinking as a way of being and thinking, emphasizing a holistic and pluralistic view of the world and making boundary judgments in different ways, as described by Reynolds and Holwell (2010). Finally, the fourth meaning is associated with the use of systems concepts. According to Patton (2010), systems thinking involves conceptualizing interventions as system change interventions rather than focusing solely on individual treatment interventions. Understanding the concept of systems thinking is crucial for conducting evaluations that aim to assess systems change interventions. The multiple meanings associated with systems thinking highlight the complexity of these types of interventions and the need for evaluators to adopt a holistic perspective that considers the whole system rather than individual components in isolation. This approach is particularly important when evaluating systems change interventions, which involve multiple stakeholders and factors that interact dynamically, making it challenging to identify and assess causality.
New Approaches and Methods to Evaluate Transformations. Meyer (2020) emphasizes the need to recognize that transforming non-sustainable systems into sustainable systems is a long-term process with interrelated dynamics across various systems. He notes that scientific research tends to embed long-term thinking more than the decision-making processes of government, private companies, and civil society organizations. Conversely, decision-makers often remain focused on local challenges, while scientists explore broader generalizations without downscaling their findings. Another perspective highlights the importance of addressing power differences in interventions and evaluation processes.

Magro et al. (2021) suggest that to address this challenge, evaluators must understand first what transformation means, its context, and its processes. Evaluating transformational change involves understanding the social, political, and economic factors that contributed to it and must consider criteria such as equity, inclusion, and sustainability instead of traditional criteria like relevance, effectiveness, efficiency, and efficacy. Consequently, new approaches and methods will be necessary. In the Global South, there have been efforts to develop tailored monitoring and evaluation approaches and methods that account for the intersections of gender, culture, and rights. For instance, the "Made in Africa Evaluation" (Chilisa, 2015) encourages data collection methods such as storytelling, folklore, music, dance, oral traditions, and African languages to challenge prevalent evaluation practices that are not culturally responsive or diverse (Sibanda & Ofir, 2021).

---

1 “The term “Global South” refers broadly to the regions of Latin America, Asia, Africa, and Oceania. It is one of a family of terms, including “Third World” and “Periphery,” that denote regions outside Europe and North America, mostly (though not all) low-income and often politically or culturally marginalized. The use of the phrase Global South marks a shift from a central focus on development or cultural difference toward an emphasis on geopolitical relations of power” (Sassen, 2013, p. 12).
In summary, the transformation of non-sustainable systems into sustainable ones is a long-term and interrelated process that requires new approaches and methods to evaluate its impact. Adopting criteria such as equity, inclusion, and sustainability and understanding the social, political, and economic factors that contributed to the transformation are essential components of the evaluation process. Furthermore, culturally responsive and diverse approaches to evaluation can facilitate a more comprehensive understanding of transformational change.

**Inclusion and Diversity.** The importance of being responsive to local contexts where transformations and impacts on people’s lives and the environment occur is a fundamental aspect of the "no one left behind" slogan embedded in Agenda 2030. In addition, the Prague Declaration (2019) emphasizes the importance of engaging and participatory evaluations to support the transformations needed to implement the Sustainable Development Goals (SDGs) (International Development Evaluation Association (IDEAS), n.d.). To achieve these goals, approaches like culturally responsive evaluation and transformative evaluation are essential. The transformative evaluation approach, developed by Mertens (2008), intentionally addresses power differences to include the voices of all stakeholders, including the most marginalized. The ultimate goal of the transformative evaluation approach is to support the development of culturally responsive interventions that promote human rights and achieve social, economic, and environmental justice.

As an example of the transformative evaluation approach, Widjaningsih and Mertens (2019) conducted an evaluation of the SDGs in West Java, Indonesia. The research team incorporated the viewpoints of different parts of society to understand the context of the need for transformational change, allowing marginalized groups such as people from rural areas, women, youth, and people with disabilities to bring their voices to the process. By hearing from different
stakeholders, the team could highlight the contrast between the perspectives of people in power and those who were not. The research team employed a mixed methods design to gather information, and farmers and youth, who are typically marginalized in decision-making processes, proposed solutions that could address social, economic, and environmental injustices.

This session addressed the key challenges faced in evaluating systems change and its implications for evaluation practice. Bamberger et al. (2016) propose a comprehensive framework that emphasizes the importance of acknowledging complexity in evaluations by considering various factors: the initiative being evaluated, the contextual environment, the stakeholders involved and their interactions, the nature of change and causality, and the evaluation process itself. In line with this, Gates et al. (2021) emphasize the necessity for evaluators to adopt a holistic perspective that encompasses the entire system rather than focusing solely on individual components. Evaluating transformations requires an understanding of the social, political, and economic factors contributing to the transformation, while also incorporating criteria such as equity, inclusion, and sustainability, and adopting new approaches and methods accordingly. Given the challenges for evaluation practice, the next session will discuss the implications for the professionals involved in evaluation.

Implications for Evaluators

Evaluation is a multifaceted field that encompasses both objective knowledge production and commissioned services that involve trade-offs and methodological compromises. It is an area of work that originates from a variety of demands. Nielsen et al. (2018) state that formal evaluations are largely shaped by the selection process that determines the evaluations’ scope, design, methodologies, and deliverables. Organizational factors such as values and beliefs, leadership style, budget, timeline, and data constraints influence these decisions. Therefore,
evaluations of global systems change initiatives, such as Agenda 2030, are naturally dependent on demand from the market side. While not all organizations have embraced the values embedded in Agenda 2030, global issues such as climate change, inequalities, and social injustice affect every initiative. As a result, there are increasing calls for all non-governmental organizations (NGOs) and programs at all levels to incorporate attention to and contribute toward global sustainability and equity (Patton, 2019).

Being a transformative evaluator entails adopting a transformative paradigm. Paradigms represent points of attention that guide thinking and behavior, providing a belief system that facilitates action (Patton, 2019). The transformative paradigm has implications in evaluators’ values and ethics (axiological assumption); in their understanding of the nature of reality (ontological assumption); in their relationship with stakeholders and in how knowledge is built (epistemological assumption); and in systematic inquiry (methodological assumptions) (Guba & Lincoln, 1989). Widianingsih and Mertens (2019), extends evaluators’ action focus to contribute to the public good, by broadening their interest toward medium and long-term results and unexpected consequences of interventions and investigating the root causes of the social problems that the interventions are designed to address (Rodriguez-Bilella et. al, 2021).

Axiological Assumption. The axiological assumption in evaluation is the underlying values, beliefs, and ethical considerations that guide evaluators in their work. It recognizes that evaluators bring their own values and biases to the evaluation process, which can influence the evaluation design, data collection, analysis, and interpretation (Mertens, 2008). Thus, it is important for evaluators to reflect on their values and biases and to be transparent about them to ensure that their evaluations are fair and ethical. Additionally, the axiological assumption recognizes that evaluation is not a value-free or neutral process but rather involves value
judgments. Evaluators must consider the ethical implications of their evaluations and strive to promote social justice, equity, and fairness. In this sense, approaches such as culturally responsive evaluation, mixed methods, and participatory methods are approaches that transformative evaluators should consider. A transformative evaluator must always be attentive to human rights and social justice issues in the evaluation process.

**Ontological Assumption.** The ontological assumption in evaluation refers to the understanding that reality is constructed and subjective, and that different stakeholders may have different perspectives on what constitutes reality. This assumption recognizes that evaluators cannot claim to have an objective or neutral view of reality but must instead acknowledge the multiple perspectives and realities that exist. Evaluators must also recognize that their interpretations of reality are shaped by their own experiences, biases, and cultural backgrounds, and that these interpretations may differ from those of other stakeholders (Mertens, 2008). Thus, it is important for evaluators to engage in dialogue with stakeholders to better understand their perspectives and to incorporate these perspectives into the evaluation process. Another ontological aspect for transformative evaluators is the importance of understanding the complexity of reality, which implies acknowledging different contextual factors and the nature of evaluands, the characteristics of the evaluands, the institutions and stakeholders involved in the initiative and how they interact, and the nature of causality and processes of change (Bamberger et al., 2016). The ontological assumption in evaluation highlights the importance of reflexivity and cultural sensitivity in evaluation practice.

**Epistemological Assumption.** The epistemological assumption in evaluation refers to the nature of knowledge, what counts as valid evidence and how it is acquired. This assumption concerns whether knowledge is objective or subjective, whether it is based on empirical evidence
or personal experience, and whether it is context-dependent or context-free. In evaluation, the epistemological assumption influences the choice of evaluation methods, and the way data are collected, analyzed, and interpreted. Some evaluation approaches assume that knowledge is objective and can be obtained through scientific methods, while others acknowledge the subjective nature of knowledge and the importance of understanding different perspectives and contexts. (Guba & Lincoln, 1989; Mertens, 2008). The choice of epistemological assumption in evaluation has implications for the validity and reliability of evaluation findings, as well as for the ethical considerations involved in evaluation practice.

**Methodological Assumptions.** Methodological assumptions are based on the belief that there are appropriate methods and techniques for obtaining valid and reliable information that can be used to make judgments and decisions about the worth and effectiveness of interventions. This assumption recognizes the importance of rigorous and systematic inquiry and emphasizes the need for evaluators to select appropriate methods and techniques based on the evaluation questions, the program’s context, and the available resources. In alignment with the three previous assumptions, it also acknowledges the importance of using multiple approaches, methods, and techniques, as well as different theories in evaluation. So, transformative evaluators are more inclined to use mixed methods, participatory and culturally responsive approaches, and identify the most important groups to include (Mertens, 2008).

These different assumptions make it possible to identify the competencies and skills needed for evaluators to evaluate systems changes. Evaluators must be able to identify the values and beliefs adopted to design the evaluation initiative (or system) and the purpose of the evaluation (axiology). Then, based on this identification, understand what counts as valid
evidence (ontology) and how the knowledge and the sense-making are intended to be built (epistemology).

**Systems Thinking.** To effectively evaluate changes in systems, evaluators must possess the ability to think dynamically about systems. They should be world-savvy evaluators who demonstrate global competence and possess knowledge in addressing world systems issues through evaluation (Patton, 2019). Evaluators must recognize the interconnectivity between projects and programs and the broader systems they operate within. Therefore, they must adopt a holistic approach that focuses on the entire system instead of its individual components. Additionally, evaluators should utilize system concepts, theories, methods, approaches, and tools to achieve a comprehensive understanding of the system.

**Complexity responsiveness.** Transformative evaluators need to be able to acknowledge that any initiative is subject to a range of contextual factors that can vary over time. They should also recognize that linear tools, such as logic models and theories of change, may not be suitable for fully comprehending the nature and characteristics of an initiative. Additionally, it is critical for evaluators to develop the ability to map stakeholders and identify those whose participation in the evaluation process is essential.

**Reality as a Social Construct.** Evaluation processes typically involve a variety of stakeholders, each with their own unique perspectives on initiatives design, implementation, and outcomes. Transformative evaluations must be sensitive to the fact that some stakeholders occupy privileged positions while others are excluded, particularly minority groups such as people of color, youth, indigenous peoples, LGBTQ+ individuals, and women. Evaluators must be aware of this and take steps to ensure that all stakeholders are given a voice and that their perspectives are considered.
Ability to Listen and Engage Stakeholders. Transformative evaluators recognize the significance of cultural aspects, actively listen to all parties, and engage with stakeholders, including those who are typically marginalized and whose voices often go unheard. They also adopt appropriate language to facilitate effective communication. By doing so, they establish trusting relationships with stakeholders. In order to build these relationships, evaluators must possess strong communication and interpersonal skills, and utilize facilitation techniques during the evaluation process.

Mastery of Approaches and Methods. Transformative evaluators employ a variety of approaches and methods, which Patton (2019) refers to as "methodological bricolage". This bricolage idea requires evaluators to utilize their creativity and eclecticism in selecting approaches and methods that are appropriate for different contexts and evaluation questions. As a result, evaluators should possess proficiency in a range of methodologies.

Youth Participation. The evaluation community must assume the commitment to engage YEEs in evaluation processes. Evaluating systems changes requires a range of complementary competencies and skills that impact both experienced practitioners and novice evaluators. Agenda 2030 highlighted the need for increased participation of youth in evaluation processes, both as beneficiaries and as members of evaluation teams. Montrosse-Moorhead et al. (2019) note that despite being relevant stakeholders, youth have been excluded from evaluation processes. They propose a youth-participatory evaluation approach that empowers young people and recognizes their unique perspectives, enabling them to play a more active role in all phases of the evaluation process. This approach moves evaluation away from being "done to young people" towards evaluation "with or by young people" (p. 40). In 2012, the Evalpartners
initiative\(^2\) proposed a Global Evaluation Agenda for the period of 2016-2020 (EvalAgenda), which included a range of components to support the professionalization of evaluation. One of the strategies proposed was to advocate for academic and training institutions to develop relevant courses in evaluation methodologies and approaches that target young and emerging evaluators (YEEs) and youth-inclusive evaluation methodologies (Evalpartners, 2016). As a result, there has been an increase in the number of novice evaluators engaged in professional associations worldwide, leading to the creation of the EvalYouth Global Network. This network supports YEEs, especially young women, in becoming technically sound, experienced, and well-networked professionals who contribute to evaluation capacity at national, regional, and international levels (Evalpartners, 2016). The above-mentioned aspects present a multitude of challenges for evaluators in the future. The next session will discuss the challenges of meeting market demands for evaluations focused on systems change and acquiring the necessary competencies and skills for such evaluations.

**The Way Forward**

The first challenge is to determine when and how the market will demand evaluations that focus on systems change. The second challenge relates to how evaluators can acquire the necessary competencies and skills for conducting such evaluations. These challenges arise as commissioners increasingly demand advanced competencies and skills for complex evaluation services. For instance, funders such as the Green Climate Fund (GCF), the Global Environmental Facility (GEF), and the Climate Investment Funds (CIF) require evaluations that assess how the work they support can achieve change. Evaluation of this kind requires a combination of

---

\(^2\) Evalpartners initiative is a partnership between the International Organization for Cooperation in Evaluation (IOCE), United Nations agencies, and other donors.
traditional evaluative skills, innovative ones, a keen sense of stakeholder facilitation, a learning orientation, and a solid grasp of systems thinking concepts and technical issues related to climate change. This range of skills is unusual, and few individuals or organizations possess them (Uitto et al., 2019).

The concept of evaluation supporting transformational change has been discussed for several years, leading international organizations to develop guidance to operationalize it. For example, the Independent Evaluation Group of the World Bank developed a framework for supporting transformational change for poverty reduction and shared prosperity in 2016, while the Global Environment Facility developed innovative evaluation methods to address climate change and other environmental issues in 2018 (Patton, 2019). One way to bridge this gap is to introduce this complementary knowledge into current evaluation education program curricula. Another strategy is to systematically research and disseminate new cases of this type of evaluation through peer-reviewed journals and conferences promoted by VOPEs. The following section will outline the research objectives and structure of the present study.

The Current Study

The main objective of the present research is to address the challenges faced by formal evaluation education programs in response to the rapidly changing world of evaluation. The study comprises two qualitative research studies aimed at exploring the perspectives of international evaluation experts and young and emerging evaluators (YEEs). The first study investigates the viewpoints of international evaluation experts regarding the role of evaluation in supporting systems changes and the knowledge, skills, and approaches that evaluation education programs should incorporate to prepare professionals to tackle these challenges. The second study focuses on the perspectives of YEEs and their expectations and needs with respect to
evaluation education. The first two studies aim to analyze the perspectives of these two groups in terms of the most important topics, approaches, and methods that should be included in evaluation education programs. A third study will map the current offerings of formal evaluation education programs and compare them with the topics that these two groups have identified as the most relevant.

The reviewed literature indicates that while there exist studies and scholarly discourse pertaining to evaluation education, there is a lack of empirical investigations involving consultations with experienced experts and the intended audience regarding their viewpoints on the constituents of an exemplary postgraduate program for evaluators. This raises inquiries regarding the authority responsible for determining the program's curriculum, the methods employed to equip students, and the specific career paths being targeted. Furthermore, the unprecedented inclusion of young and emerging evaluators (YEEs) in the research process, which considers their barriers and needs, would add a unique and forward-looking dimension to this study.

This research will provide valuable insights into the alignment between academic programs and the evolving needs and expectations of the field of evaluation. By assessing the inclusion of essential topics, approaches, and methods, the study can identify any gaps or discrepancies that may exist in current program curricula. Secondly, such a study can inform program development and improvement efforts by identifying areas where enhancements are needed to ensure graduates possess the necessary knowledge and skills for effective evaluation practice. Additionally, the study can contribute to advancing the field of evaluation education by providing evidence-based recommendations for curriculum design and content. Ultimately, the
study's findings can support the overall professionalization and advancement of evaluation as a discipline.

**Research Questions**

Intending to inform how formal evaluation education programs respond to the field’s challenges in a context of a world in transformation, the proposed study aims to address the following questions.

1. What are the international evaluation experts’ perspectives regarding the idea that evaluation has a role in supporting systemic changes?
2. When looking for opportunities in evaluation education, what are the main barriers YEEs face?
3. What are the YEEs main interests in terms of themes, approaches, and methods?
4. What should be included in education programs to prepare professionals involved in systemic changes evaluations?
   a. Which topics, approaches, and methods are likely to gain more importance for the evaluation practice in the coming years?
   b. Should practicum be required as part of the curriculum?
5. To what extent do existing evaluation education programs address relevant topics, approaches, and methods to prepare evaluation professionals for the field’s challenges?
6. To what extent do existing evaluation education programs address YEEs needs and interests?

The forthcoming chapter expounds upon the methodologies implemented to address the research questions posited in this study. Following a presentation of the
research design and its overview, a comprehensive description of the phases and corresponding studies' methodologies will be explicated in detail. This will encompass a discussion of how participants were selected and included the elucidation on the data collection instruments and procedures employed, and a delineation of the plan for analysis.
Chapter 3: Methods

Research Design and Overview

To address the research questions mentioned in the previous chapter, the study used a sequential multi-phased mixed-methods design, conducted in two phases (Figure 1). Phase 1 consisted of two qualitative studies. Study 1 explored international evaluation experts’ perspectives regarding the evaluation role in supporting systems change, as well as the critical topics, approaches, methods, and other characteristics that should be present in formal evaluation education programs to prepare evaluators for future challenges in a world in transformation. It consisted of gathering qualitative data through semi-structured interviews. The results of Study 1 informed the development of Study 2 which aimed to collect perspectives on the role of evaluation in supporting systems changes and interests and needs regarding evaluation education from a group of professionals representing the field’s future workforce: the young and emerging evaluators (YEEs). In Phase 2, a mixed-methods study was conducted to compile a list of formal postgraduate evaluation education programs (EEP) provided by academic institutions globally and to assess to what extent these programs cover the necessary content, topics, approaches, and methods to equip evaluation professionals with the skills needed to tackle challenges in an ever-changing world. Additionally, this study aimed to assess the extent to which the programs cater to the needs and interests of young and emerging evaluators (YEEs) regarding evaluation education. Specifically, the study examined the descriptions and curricula of university-based graduate evaluation education programs across various degrees such as diplomas, certificates, master’s, and doctoral degrees, as well as other operational characteristics such as location, modality, and disciplinary focus.
Phase 1

Phase 1 included the following two studies: Study 1, qualitative semi-structured interviews with international evaluation experts and Study 2, a focus group with young and emerging evaluators (YEEs). This phase had the purpose of exploring the two groups’ perspectives on the role of evaluation in supporting systems changes and the critical topics, approaches, and methods that should be incorporated into formal evaluation education programs to prepare evaluators for future challenges in a transforming world. In addition, Study 2 explored YEEs’ challenges and interests pertaining to evaluation education.

Study 1. Study 1 was a qualitative study using a semi-structured interviews.

Participants. In Study 1, 19 evaluation specialists from different backgrounds, roles and continents were interviewed. The selection process used a purposive reputational sample snowball method (Creswell & Poth, 2018). The initial five subjects invited (Appendix B: Draft Email Invitation for Interview – First Subjects) were leaders of international organizations that
share the mission to advance evaluation as a profession across the globe: the International Development Evaluation Association (IDEAS), the International Evaluation Academy (IEAc), the Global Evaluation Initiative (GEI), and the International Organization for Cooperation in Evaluation (IOCE). At the end of the interviews, those five representatives were asked to nominate three to four other experienced experts who have contributed to the debate on the importance and role of evaluation in improving initiatives, policies, and change systems over the last years. The nominated experts should represent the following groups: leaders of formal evaluation education programs; researchers on evaluation; authors of books and articles that discuss evaluation as a professional field; leaders or former leaders of Voluntary Organizations for Professional Evaluation (VOPEs); and multicultural evaluation practitioners and commissioners. A list of 24 potential participants was generated based on the first group of interviewees’ recommendations. From the list of names generated, the PI used purposive sampling to select participants who fulfilled the criteria. All 24 of the specialists were invited to participate in an interview (Appendix C: Draft Email Invitation from Interview – Referred Subjects). The PI received a positive answer from 19 (79.17%) of the 24 invitees. For those who accepted the invitation, the interviews were scheduled via email. Prior to the interview, participants received a reminder about the scheduled interview and an informed consent form providing them with additional information about the study (Appendix D: Draft Email Reminder for Interview, Appendix E: Draft Interview Consent Form).

Most participants (n = 11, 57.90%) were women, and white (n = 15, 78.95%). They originated from six regions: Africa (3); Asia (2); Europe (3); Latin America and the Caribbean (5); North America (4); and Oceania (2). In terms of the most advanced educational degree, most interviewees (n = 12, 63.16%) had a doctoral degree, and seven of them (36.84%) had a master’s
Furthermore, over half of the interviewees (n = 10, 52.63%) reported having received specific training or education in evaluation. Regarding work experience, all participants had more than 14 years of experience. The interviewees’ professional backgrounds varied from teaching activities, research, consulting, and evaluation commissioning. In addition, a large proportion of the participants (n = 13, 68.42%) had a record of volunteering in professional organizations or initiatives that aimed to advance evaluation as a profession in their countries, regions, or even internationally.

Instrument. Study 1 gathered qualitative data from international evaluation experts through semi-structured interviews. The semi-structured interviews were divided into three primary sections, as outlined in Appendix A: Interview Protocol. The first section aimed to establish rapport with the interviewees, allowing them to feel more comfortable throughout the interview. Given the diversity of the sample, this section aimed to gain insight into their professional background in evaluation and their main contributions to the field. The second section focused on the interviewees’ perspectives on the role of evaluation in an era of transformation. Participants were asked to provide their viewpoint on systemic or transformative change, as well as the role that evaluation can play in contributing to such change. Then, given the transformative change context, the following question probed their perspective on topics, approaches, and methods likely to gain importance for evaluation practice in the coming years. Participants were also asked whether they see these trends becoming evident in the field and how they have impacted their professional practice. The section ended with a question about the challenges ahead for evaluation as a field in an era of transformative change. In the third and final section, respondents were asked to share strategies that have been employed and any barriers or obstacles that may impact the effectiveness of such strategies, given the challenges
presented. The interviewees were also asked about the most important implications of the discussed changes for the formal preparation of new evaluators, particularly in university settings. Probes were used when necessary, during the interviews, which concluded with an expression of gratitude from the PI.

The interview protocol was subject to discussion with the PI’s advisors and piloted with the first interviewee, leading to subsequent refinement of the questions and improved interaction with the interviewer. Additionally, the CGU Institutional Review Board reviewed and approved the interview protocol, ensuring that relevant ethical aspects of the research were appropriately addressed. The refinement and approval of the research instrument were crucial steps to ensure quality assurance in this study.

**Data Collection Procedure.** The semi-structured interviews were conducted and recorded via Zoom Platform and lasted 30 to 45 minutes. The PI followed the interview protocol but was open to gathering additional information relevant to the study. The interviewer started by reading the informed consent information to the respondents. Once consent was given, the interviewer asked to voice-record the interview, putting the camera off to preserve the interviewees’ images. All the participants agreed to have the interview recorded. Therefore, all the interviews were recorded on ‘Zoom Platform’ and a separate recording device. Data from the interviews were transcribed with the support of the application ‘otter.ai’ and organized into electronic folders protected by a password.

The sample size of this study \((n = 19)\) could be considered small; however, the determination of an ideal or fixed sample size for ensuring quality in qualitative research is not possible. The required number of interviews in qualitative studies depends on various factors, including research questions, scope of study, intended use of findings, and available resources.
for the study (Patton, 2015). In this study, the researcher chose to focus on obtaining a broader range of perspectives and a more in-depth discussion from a smaller number of participants. To achieve this goal, the researcher made sure the sample was formed by a group of seasoned evaluation professionals who had diverse educational backgrounds, origins, career paths, and roles in their professional lives. This approach was consistent with the focus of qualitative research on obtaining detailed insights into a particular subject rather than generalizing findings. The researcher also realized a point of data saturation in the last interviewees, as they did not provide new insights compared to previous participants.

**Plan for Analysis.** The qualitative data analysis proceeded according to the following process. First, the PI proceeded with a data cleaning. Through the cleaning process, all the transcriptions were verified by checking them against the audio recordings of the interviews, aiming to identify errors or inconsistencies. During some of the interviews, the conversation deviated from the main focus of the protocol due to their reflective nature on the participants' careers, which are both exciting and challenging in the field of evaluation. The researcher acknowledged that these digressions were not related to the research objectives and ended up being personal anecdotes. Therefore, for the sake of data quality, the cleaning process removed these extraneous parts, allowing for a better focus on the data relevant to the study. The cleaning process also intended to remove words and expressions commonly used in informal speech, such as asides and repeated points. Furthermore, the cleaning process allowed for a first read-through of the data and helped in consolidating the responses for each of the interview questions.

Following the data cleaning process, a preliminary set of seven codes was established, as presented in Table 1. These codes were derived from the structure of the interview script and used as an initial framework for the subsequent analysis.
### Table 1

*Initial List of Codes for Interviews Analysis*

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>i) Background</td>
<td>Background of the interviewees, highlighting their primary contribution to the evaluation field in their contexts</td>
</tr>
<tr>
<td>ii) Transformation and evaluation role</td>
<td>What terms such as transformation, transformational change, and systems change mean for the participants and their vision about the role of evaluation in contributing to transformations</td>
</tr>
<tr>
<td>iii) Implications for the practice</td>
<td>Implications of the role of evaluation in contributing to transformations for evaluation practice and what topics, approaches and methods are likely to gain importance in the future</td>
</tr>
<tr>
<td>iv) Challenges at a global level in an era of transformational changes</td>
<td>Challenges at a global level for the practice of evaluation in an era of transformational change</td>
</tr>
<tr>
<td>v) Strategies that have been employed given the challenges</td>
<td>Given the challenges discussed, strategies that have been observed in the evaluation community to transform evaluation</td>
</tr>
<tr>
<td>vi) Barriers and Obstacles</td>
<td>What are some barriers or obstacles that may impact on the effectiveness of such strategies</td>
</tr>
<tr>
<td>vii) Implications for evaluation education</td>
<td>How evaluation education should change to better prepare evaluation professionals given the implications for the practice and the challenges discussed</td>
</tr>
</tbody>
</table>

The researcher used ‘MaxQDA’, a software program designed for computer-assisted qualitative and mixed methods data, text, and multimedia analysis, to set up the initial code system, upload the cleaned transcripts, and attribute codes to excerpts of text through another detailed reading of the transcripts. Along the coding process, additional topics, classified as sub-codes, emerged and were added to the primary codes, resulting in a total of 88 sub-codes.
associated to the seven original primary codes. The code structure is provided in the Appendix K. Using MAXQDA to assist with data management and analysis helped to improve the accuracy, organization, and efficiency of the coding process.

Due to limited resources, it was not possible to have multiple coders for independent analysis of data and assess inter-coder reliability during the qualitative analysis of interviews. This poses a potential threat to the credibility of the data analysis. However, there were some characteristics of the process that can be highlighted as mitigators of these threats. Firstly, the coding system was structured based on the interview protocol, which provided coherence to the process analysis. Secondly, the extensive and multi-methods experience of the researcher as an evaluator for over 30 years, along with over 20 years of continuous study and leadership roles in national and international VOPEs, proved invaluable in comprehending the significant findings that emerged from the data and, therefore, merited reporting.

After completing the coding process, the data were exported to an ‘Excel’ spreadsheet, which facilitated sorting and grouping the excerpts by code. This process allowed for the identification of patterns, which aided in extracting relevant points to be presented in the findings section.

Study 2. Study 2 was a qualitative study using a focus group. The focus group was a closed event held online on June 2nd, 2022, in English, as part of the Glocal Evaluation Week. Participants. The focus group’s participants were all representatives of EvalYouth networks recruited through the Glocal platform, which announced all the 2022 events. The

---

3 The Global Evaluation Initiative (GEI) convenes Glocal Evaluation Week annually, during which Centers for Learning on Evaluation and Results (CLEAR Centers) facilitate knowledge sharing events worldwide. The researcher suggested the inclusion of the focus group as one such event during the 2022 Glocal Evaluation Week and obtained permission to use the data of registered participants for subsequent communication.
recruitment also relied on email requests for support from leaders in global, regional, and national chapters of the EvalYouth networks to disseminate the focus group (Appendix G: Draft Email asking for EvalYouth chapters leaders’ support to disseminate the Focus Group). In addition, personal messages to EvalYouth networks leadership were sent by email. Through the EvalYouth Global portal, the researcher gathered the seven regional and 19 national chapters’ contacts, websites, and social media channels. Then, in addition to email messages, a message with the invitation and the flyer (Appendix J: Flyer for Dissemination of the Focus Group in EvalYouth Chapters’ Social Media Channels) was posted on all social media channels listed on the EvalYouth Global portal, including chapters’ Facebook fan pages, Twitter pages, and LinkedIn groups. The invitation message informed viewers that the maximum number of participants would be 15, and YEEs interested in participating were directed to fill out a registration form provided by the CLEAR Center, organizer of the 2022 Glocal. Interested participants filled out a registration form provided by the CLEAR Center, the organizer of the 2022 Glocal Evaluation week, and personal emails were sent to registered participants with the consent form. By signing the consent form, the participants declared that they fulfilled the participation eligibility criteria, that is, they were under 35 years of age, or had less than five years of professional experience and, within the past five years or less, had expressed an interest in becoming a professional evaluator.

From the contact information provided by the registration records, accessible through a PI’s Glocal portal exclusive login and password, personal emails were sent to the registered YEEs thanking them for their interest in participating and sending them the consent form (Appendix H: Focus Group Consent Form), which informed them that the session would be
recorded. Prior to the Focus Group, registered YEEs received a reminder about the scheduled meeting (Appendix I: Draft Email Reminder for Focus Group Participants).

Study 2 counted on the participation of 11 YEEs from the 20 initially registered. Of the 11 participants, seven were female, and four were male. Regarding country of residence, four resided in Latin America, four in Africa, two in North America, and one in Asia. Regarding their institutional affiliation, four participants declared they were professionally dedicated to VOPEs; three worked for non-governmental organizations; two were affiliated with academic institutions; one worked for a multilateral development organization; and one worked for a private company. Finally, participants were asked to identify their primary role in the evaluation field. Of the 11 participants, eight declared themselves as evaluation practitioners, two were students, and one was a research assistant.

The purpose of the recruiting was to have a group as diverse as possible, which was facilitated by the capillarity and the presence of YEEs on social networks. As in Study 1, the researcher chose to have a broader range of perspectives and a more in-depth discussion from a smaller number of participants.

**Instruments.** The focus group protocol included semi-structured questions (Appendix B: Focus Group Protocol). The questions were based on the PI’s research questions and those adopted in Study 1. It included three sections. In the first session, the PI briefly presented the research context and reasoning and the focus group’s agenda to all participants. The second session started with a brief presentation of all participants. It aimed to make participants comfortable and stimulate a general discussion about what does transformational change mean to YEEs, how evaluation can support transformative change and why education is crucial to acquiring knowledge and skills. Finally, for the third part, participants were randomly assigned
to three discussions group and had 40 minutes to discuss three questions: i) What challenges do YEEs face to access education opportunities; ii) What should be the priorities for evaluation education programs to meet YEEs’ needs; and iii) What should evaluation education programs look like to prepare YEEs to be involved in systems change evaluations?

To improve data collection quality, EvalYouth Global leaders were previously asked to revise the focus group protocol to guarantee adequate language to support the researcher on this contextualization. The leaders recommended the use of a more contextual language during the discussions because less experienced evaluators formed the target audience.

**Procedures.** The focus group lasted two hours and was conducted and recorded via ‘Zoom Platform’. At the end of the second session, the PI asked three volunteers to help facilitate and voice-record the group discussions. This last session ended with a plenary discussion, which was also recorded. Data from the focus group were transcribed with the support of the application ‘otter.ai’ and organized into electronic folders protected by a password.

**Plan for Analysis.** The qualitative data analysis proceeded according to the following process. First, the PI conducted a data-cleaning process. Through the cleaning process, all the transcriptions were verified by checking them against the audio recordings of the focus group’s main room, breakout rooms, and plenary session, aiming to identify errors or inconsistencies. The cleaning process also intended to remove parts that were not related to the nature of the interview or words and expressions commonly used in informal speech, such as asides and repeated points. The cleaning process allowed for a first read-through of the data and helped in consolidating the responses for each of the interview questions. Then, the researcher defined a list of four initial codes based on the structure of the focus group protocol, presented in Table 2.
Table 2

Initial List of Codes for Focus Group Analysis

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>i) Transformation</td>
<td>Meaning of transformational change to YEEs</td>
</tr>
<tr>
<td>ii) Challenges to access evaluation education opportunities</td>
<td>Challenges YEEs face to access evaluation education opportunities, such as personal challenges, barriers created by organizations they work for, evaluation education offer, affordability, and curriculum appropriateness</td>
</tr>
<tr>
<td>iii) evaluation education priorities</td>
<td>What should be the priorities for evaluation education programs to meet YEEs’ needs</td>
</tr>
<tr>
<td>iv) evaluation education characteristics to prepare YEEs for transformative change</td>
<td>Which topics, approaches, and methods are likely to gain more importance in the evaluation practice in the coming years</td>
</tr>
</tbody>
</table>

The PI used ‘MaxQDA’, a software program designed for computer-assisted qualitative and mixed methods data, text and multimedia analysis, to set up the code system, upload the transcripts, and attribute codes to excerpts of text through another detailed reading of the transcripts. Along the coding process, sub-codes emerged from the readings and were added to the primary codes. A total of five sub-codes were created under group (iv): topics, approaches, methods, engagement in communities of practice; and requirement of practice. The absence of additional subcodes for the remaining primary codes can be attributed to the fact that the focus group yielded more targeted discussions concerning the four themes, as opposed to the interviews.

Phase 2

Study 3. The second phase of this research aimed to compile a list of formal postgraduate evaluation education programs (EEP) provided by academic institutions globally and to assess to
what extent these programs cover the content, topics, approaches, and methods identified by participants as needed to equip evaluation professionals with the skills to tackle challenges in a world in transformation. Additionally, this study aimed to assess the extent to which the programs cater to the needs and interests of young and emerging evaluators (YEEs) regarding evaluation education. Specifically, the study examined the descriptions and curricula of university-based graduate evaluation education programs across various degrees such as diplomas, certificates, master’s, and doctoral degrees, as well as other operational characteristics such as location, modality, and disciplinary focus.

**Sample.** Study 3, a mixed-methods study, utilized a combination of archival and website search techniques to collect data on postgraduate evaluation education programs worldwide. To do this, the PI employed an approach utilized by LaVelle (2014) and LaVelle and Donaldson (2015), which involved obtaining an initial sample frame from previously published articles and directories of programs, then supplementing it with information from other sources. Firstly, the researcher gathered data from directories and lists of programs from recent articles cited in King and Ayoo’s (2020) study, as well as other directories such as Latin America and the Caribbean (Ghiano et al., 2017) and Africa (Basheka & Byamugisha, 2015; Wildschut & Silubonde, 2020). The initial list was further supplemented with listings from the American Evaluation Association’s education opportunities list (AEA, 2022), the Canadian Consortium of Universities for Evaluation Education (CUEE) post-secondary courses list (Evaluation Education, 2023), the Réseau Francophone de l’Évaluation, and the European Evaluation Society’s education opportunities list (European Evaluation Society, n.d.). After compiling data from these sources, the researcher compared the list of countries identified as offering EEP to the complete list of countries and identified countries that did not offer EEP. Then, the researcher crossed the names
of those countries and the area of influence of regional VOPEs and CLEAR Centers. Electronic messages were then exchanged between the researcher and the leadership of those regional VOPEs and CLEAR Centers to consult them about the existence of academic programs in their region, aiming to fill any gaps in the data. This communication included contacting the Asia-Pacific Evaluation Association (APEA), African Evaluation Association (AfrEA), Evaluators Network of the Middle East and North Africa (EvalMENA), CLEAR Center Francophone Africa, and CLEAR Center South Asia. To gather contact information from the leaders of the institutions, the researcher utilized the IOCE VOPEs directory and the CLEAR Center websites. After the second set of data was provided by these organizations was incorporated into the initial list, the researcher concluded that there were no additional sources of reliable information and finalized the preliminary list of programs, which consisted of 407 programs.

After finalizing the initial list of programs, the PI undertook a rigorous data supplementation, validation, and cleaning process. This process of data supplementation was deemed necessary because the initial list of programs lacked crucial information required for the research, such as program objectives, learning outcomes, curricula, cost, time to degree and language. These pieces of information were imperative to address the research questions effectively. Due to the publication directories being dated over four years ago and the disruptive impact of the COVID-19 pandemic on institutional changes and discontinuity, the researcher deemed it necessary to conduct a validation process. The data supplementation and validation process involved an exhaustive examination of each university’s official website. To identify the university program’s weblinks, the researcher queried the World Wide Web through the Google Chrome page browser and accessed the links housed under the official websites of the
universities. Data collection and validation were conducted between April 2022 and February 2023. The dataset followed the structure detailed below:

- Region/Sub-region/Continent/Country\(^4\)
- Classification as Global South country\(^5\)
- Name of College/School/Department
- Type of Program (master’s, doctorate, certificate, diploma, etc.)
- Title of Program
- Method (presential, online, blended, hybrid)
- Website link
- Contact information
- Number of U.S. credit hours
- Evaluation core/required courses/credits
- Elective courses/credits
- Other requirements
- Tuition for international students
- Time to completion
- Language

The cleaning process consisted of eliminating programs not found on the Universities’ websites or programs that did not fulfill the inclusion criteria. As mentioned before, the researcher adopted the same operational definition of "program" used by Altshuld et al. (1994), cited by King and Ayoo (2020): "multiple courses, seminars, practica, offerings, and so on designed to teach evaluation principles and concepts" (p.3). Therefore, the final list of programs included university-based graduate degrees (master’s or doctoral), diplomas, and certificates offered by universities or teaching institutions, of any modality, in any language, and offering at

\(^4\) According to [https://unstats.un.org/unsd/methodology/m49/](https://unstats.un.org/unsd/methodology/m49/)
\(^5\) According to the [United Nations’ Finance Center for South-South Cooperation](https://unstats.un.org/unsd/methodology/m49/)
least two courses with the word "evaluation" in the title. Therefore, this study did not include professional development programs and settings that offer only one or two evaluation courses or single workshops. After the cleaning and validation process, 271 programs from 39 countries and six continents were included in the final list.

The primary characteristics of the sample, categorized by continent, are presented in Table 3. Alongside the classification by continent, the researcher incorporated another classification showcasing programs in countries forming part of the Global South, following the classification of the Finance Center for South-South Cooperation⁶.

Table 3

Summary of Evaluation Education Programs Across the World (N=271)

<table>
<thead>
<tr>
<th>Continent</th>
<th>North America</th>
<th>South America</th>
<th>Europe</th>
<th>Africa</th>
<th>Asia</th>
<th>Oceania</th>
<th>Total</th>
<th>Global South</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>1) Access</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Countries with evaluation degrees</td>
<td>40</td>
<td>58</td>
<td>18</td>
<td>42.9</td>
<td>17.5</td>
<td>8.7</td>
<td>4.3</td>
<td>1</td>
</tr>
<tr>
<td>Total of Countries by Continent</td>
<td>40</td>
<td>58</td>
<td>18</td>
<td>42.9</td>
<td>17.5</td>
<td>8.7</td>
<td>4.3</td>
<td>1</td>
</tr>
<tr>
<td>2) Type of Degree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diploma</td>
<td>14</td>
<td>14</td>
<td>15</td>
<td>15</td>
<td>10</td>
<td>4</td>
<td>10</td>
<td>13.7</td>
</tr>
<tr>
<td>Certificate</td>
<td>52</td>
<td>50</td>
<td>58</td>
<td>51</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>100.0</td>
</tr>
<tr>
<td>Master’s</td>
<td>60</td>
<td>45</td>
<td>59</td>
<td>76</td>
<td>49</td>
<td>43</td>
<td>49</td>
<td>100.0</td>
</tr>
<tr>
<td>Doctoral</td>
<td>40</td>
<td>22</td>
<td>34</td>
<td>12.1</td>
<td>33</td>
<td>100.0</td>
<td>33</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>175</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>3) Thematic Focus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalist</td>
<td>36</td>
<td>20</td>
<td>20</td>
<td>4.3</td>
<td>36</td>
<td>20</td>
<td>36</td>
<td>100.0</td>
</tr>
<tr>
<td>Specialized</td>
<td>139</td>
<td>74</td>
<td>28</td>
<td>80.0</td>
<td>18</td>
<td>57.6</td>
<td>57.6</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>175</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>4) Modality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On Campus</td>
<td>85</td>
<td>86</td>
<td>0</td>
<td>100.0</td>
<td>40</td>
<td>0</td>
<td>40</td>
<td>50.0</td>
</tr>
<tr>
<td>Online</td>
<td>41</td>
<td>23</td>
<td>13</td>
<td>37.1</td>
<td>12</td>
<td>4.3</td>
<td>4.3</td>
<td>100.0</td>
</tr>
<tr>
<td>Mixed</td>
<td>29</td>
<td>16.6</td>
<td>14.3</td>
<td>5.0</td>
<td>15.2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>No information available</td>
<td>11</td>
<td>6.3</td>
<td>22</td>
<td>62.9</td>
<td>1</td>
<td>4.3</td>
<td>- 0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>175</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>5) Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>152</td>
<td>86.9</td>
<td>1</td>
<td>2.9</td>
<td>7</td>
<td>30.4</td>
<td>30.4</td>
<td>100.0</td>
</tr>
<tr>
<td>French</td>
<td>12</td>
<td>6.9</td>
<td>0</td>
<td>34.8</td>
<td>5</td>
<td>26.1</td>
<td>26.1</td>
<td>100.0</td>
</tr>
<tr>
<td>Spanish</td>
<td>11</td>
<td>6.3</td>
<td>22</td>
<td>62.9</td>
<td>1</td>
<td>4.3</td>
<td>- 0</td>
<td>0</td>
</tr>
<tr>
<td>Portuguese</td>
<td>0</td>
<td>- 0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other Languages</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No information available</td>
<td>0</td>
<td>- 0</td>
<td>0</td>
<td>- 0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>175</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

⁶ http://www.fc-scc.org/en/partnership_program/south_south_countries, accessed on 03.27.2023
In relation to the quality of data gathered, certain limitations ought to be acknowledged. Firstly, the researcher relied on data originating from official websites, which may be dynamic, incomplete, and subject to change. Secondly, disparities existed regarding the availability and precision of the data across various countries, which may have influenced the coding procedure. During the validation and cleaning procedure, absent data were eliminated from most of the variables coded across the 271 cases.

The data collected in this study were shared with the Global Evaluation Initiative (GEI) for the purpose of creating and publicly disseminating a comprehensive global directory of postgraduate academic programs through the "Better Evaluation" knowledge platform. The development of this website entailed the collaboration of GEI's IT team to ensure data security and protection. Additionally, a team of expert consultants, invited by the GEI, contributed their insights to optimize the presentation of the data for a diverse and global audience. Importantly, the researcher's provision of data to the GEI did not affect the process of data collecting, cleaning and analysis.

*Plan for Analysis.* From the final list obtained in the validation, a rating system was conducted to create categories for each aspect of analysis, as shown in Table 4.
Aspects of EEP Analysis and Rating System

The analysis of the data gathered allowed for the identification of topics, approaches, methods and other characteristics that should be part of evaluation education programs to better prepare professionals for the challenges related to future evaluation demands. In addition, the availability of courses was analyzed by comparing the number of programs per country between the Global South and Global North. The classification of countries as part of the Global South and Global North relied on the Finance Center for South-South Cooperation list.

The researcher analyzed each program individually and assigned codes using the coding system presented in Table 4. The attribution of codes was based on the analysis of the descriptions and names of both core and elective courses as well as other requirements.

<table>
<thead>
<tr>
<th>Aspect of Analysis</th>
<th>Categories of analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thematic Focus</td>
<td>Generalist</td>
</tr>
<tr>
<td>Core Curriculum</td>
<td>None of the 4 elements are declared or suggested</td>
</tr>
<tr>
<td>Topics</td>
<td>One topic declared or suggested</td>
</tr>
<tr>
<td>Approaches</td>
<td>Two Approaches suggested or declared</td>
</tr>
<tr>
<td>Methods</td>
<td>Three or more Approaches suggested or declared</td>
</tr>
<tr>
<td>Inclusion of New Technologies</td>
<td>Explicitly declared</td>
</tr>
<tr>
<td>Inclusion of Practicum</td>
<td>Explicitly declared</td>
</tr>
<tr>
<td>Inclusion of Project Management specific content</td>
<td>Explicitly declared</td>
</tr>
<tr>
<td>Inclusion of Communication specific content</td>
<td>Explicitly declared</td>
</tr>
<tr>
<td>Professionalization</td>
<td>Explicitly declared</td>
</tr>
<tr>
<td>Competencies</td>
<td>Explicitly declared</td>
</tr>
</tbody>
</table>

Table 4
Additionally, the researcher examined associations between variables, which were considered independent, using Pearson's chi-square test and Likelihood Ratio. Differences among groups were considered statistically significant when the \( p \)-value was less than .005. Descriptive statistics were used in other parts of the analysis.

The previously mentioned issues related to data quality potentially impacted the quality of the analysis, as the researcher may have assessed certain program characteristics based on information solely available on the websites. The analysis of topics, approaches, and methods could have been more precise if the researcher had access to the course syllabi, which are often unavailable on the websites.
Chapter 4: Results

This chapter sequentially presents the analysis results of the three studies. The first session delves into a comprehensive analysis of the interviews to explore the multifaceted aspects of evaluation and its role in contributing to transformative change. It examines how evaluation practices should evolve to drive systemic changes in various contexts. Additionally, it investigates the emerging topics, approaches and methods that are likely to gain significance in evaluation practice in the coming years. It critically examines the challenges that lie ahead for the field of evaluation in an era of transformative change, while also proposing strategies to overcome these obstacles and ensure the effectiveness of the transformation. Finally, the chapter discusses the implications of these findings for evaluation education, highlighting the need for educational programs to align with the changing landscape of evaluation and equip professionals with the necessary skills and knowledge. Through a comprehensive analysis of these interviews, this chapter provides valuable insights into the evolving nature of evaluation and its potential to drive positive and impactful change in diverse settings.

The next session provides a comprehensive exploration of the perspectives and insights of YEEs through the focus group analysis. It encompasses their understanding of transformational change, identifies emerging themes and topics in evaluation practice, discusses approaches and methods for facilitating transformative change, emphasizes the inclusion of YEEs in evaluation processes, highlights their expectations and needs regarding evaluation education, explores evaluation programs for systems change evaluations, and underscores the value of engagement in communities of practice.

The third and last session presents the findings of a comprehensive analysis of various academic evaluation education programs, focusing on presence in their curricula of key aspects,
including thematic focus, core curriculum dosage, and the offer of elective classes. Additionally, it investigates the presence of topics identified as potentially crucial for future evaluation practice, as well as approaches that are likely to gain importance in the field of evaluation amidst a rapidly transforming world. The analysis also delves into methods and new technologies employed in evaluation, the inclusion of practicum experiences and the development of soft skills, and the overall professionalization of the evaluation field. The study also explores associations between independent variables through chi-square tests of independence. By examining these dimensions, this chapter contributes to a deeper understanding of the diverse aspects and considerations involved in academic evaluation education programs.

**Study 1: Interviews with International Evaluation Experts**

Study 1 aimed to explore international evaluation experts’ perspectives regarding the evaluation role in supporting systems change, as well as the critical topics, approaches, methods, and other characteristics that should be present in formal evaluation education programs to prepare evaluators for future challenges in a world in transformation. It consisted of gathering qualitative data through semi-structured interviews. Following this, the analysis findings will be presented, utilizing a coding system consisting of two alphabetic characters, one numeric character, and three alphabetic characters to identify interviewees according to their role within the evaluation field, sequential number assigned to each interviewee, and their geographical region of origin. Table 5 provides an overview of the coding system adopted to designate the interviewees.
Table 5

Coding System Adopted to Designate the Interviewees Based on their Role within the Evaluation Domain and Geographical Region of Origin

<table>
<thead>
<tr>
<th>Code 1</th>
<th>Role within the Evaluation Domain</th>
<th>Code 2</th>
<th>Region of Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Academic / Scholar</td>
<td>AFR</td>
<td>Africa</td>
</tr>
<tr>
<td>IV</td>
<td>Leader of International VOPE</td>
<td>ASI</td>
<td>Asia</td>
</tr>
<tr>
<td>NV</td>
<td>Leader of National VOPE</td>
<td>EUR</td>
<td>Europe</td>
</tr>
<tr>
<td>RI</td>
<td>Head of Monitoring and Evaluation and Learning of a Research Institute</td>
<td>LAC</td>
<td>Latin America and the Caribbean</td>
</tr>
<tr>
<td>HP</td>
<td>Head of Evaluation Education Program</td>
<td>NAM</td>
<td>North America</td>
</tr>
<tr>
<td>PR</td>
<td>Practitioner</td>
<td>OCE</td>
<td>Oceania / Australasia</td>
</tr>
<tr>
<td>DB</td>
<td>Development Bank (funder or commissioner)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Findings

The participants in the interviews offered invaluable insights on their viewpoints regarding systems and transformative change, the potential contribution of evaluation to such change, key future trends, and the necessary adaptations to evaluation practices. Moreover, the participants also shared their perspectives on the challenges and obstacles associated with adapting evaluation to future trends and the implications for evaluation education. Throughout the interview, the participants provided compelling examples and justifications to support their opinions.

The Role of Evaluation in Contributing to Transformative Change. In this part, the researcher analyzes the interviewees’ perspectives regarding evaluation’s role in contributing to
transformative change. The evaluation community has debated in recent years the issues the planet has faced in the last few decades and the changes needed to solve these issues. Terms like systemic and transformative change have been used to refer to these changes. Before discussing the role of evaluation in contributing to transformative change, interviewees were invited to explain what these terms mean to them.

One interviewee, a leader of an international VOPE, conceptualized systemic change as an overarching, integral, structural, substantive, and interrelated change that alters existing paradigms and relates to how institutions work and interact, which reflects many of the other interviewee’s responses. One interviewee from India recognized that, in his context, people are still unsure about what transformation is about. Even though they know that some radical changes in the systems are necessary, they are not sure what those changes would be and how they can be brought into the system. He reports that this discussion is still in the primary stage in his context.

Transformative and systemic change happens holistically across different levels and dimensions and aims to bring about a profound change in the systems. From the perspective of an evaluation leader for a European research institute, sustainable development has at least four dimensions: environmental, social, economic, and governance. In his perspective, to consider that a sustained change has happened, there must have been some sort of positive change across all the dimensions, which essentially brings a system to change profoundly into something else. He illustrates this concept by saying that a systemic change is happening at the climate level, evidenced by satellite images, and that data analysts are getting extremely worried about the predicted effects of this change. One of these effects is that we will face a massive challenge regarding food security, which could trigger other challenges. Also, one of the interviewees gave
examples of how climate change is impacting local communities and affecting mostly vulnerable populations. So, transformative and systemic change depends on a multidimensional view of dimensions and levels simultaneously. A helpful way to visualize the meaning of the term ‘level’ is to think about government levels. In general, local policies decentralize national policies and if they don’t work in an articulate manner no effective change in the systems will be successful. Therefore, evaluating systems change requires teams to consider multiple dimensions and how each of these dimensions affects different levels or territories. By doing this, teams are better equipped to have a holistic view of the components that must be considered in an evaluation.

Interdisciplinarity and systemic change are mutually reinforcing concepts that can work together to solve complex problems and create a more sustainable and equitable world. From one interviewee’s perspective, changes at the system’s level imply addressing multiple parts of the system and understanding what interactions among these parts are necessary to solve social problems. She points out that a change in one part can have a catalytic effect on other parts, and even parts that are not affected by that change can affect other parts of the system. She illustrates this by saying there is a need to acknowledge that human systems impact non-human systems and vice-versa. For an interviewee from Latin America, understanding systemic interactions requires interdisciplinarity. She explains, “What we normally see is social policies that don’t talk to each other, education from one side, health from the other, for example. This lack of policy integration is not likely to promote transformations.” (IV3-LAC). So, to change a system there is a need to visualize and understand each of its parts, how they interact, and how each part and interaction contributes to change. Evaluating systems change requires teams to bring together multiple perspectives and build multiple and integrated theories of change. By bringing together
multiple perspectives, interdisciplinary teams are better equipped to identify and address the root causes of problems leading to more effective evaluation designs.

One example of paradigm, brought by other interviewees, is power relations. Systemic change pertains to changes in power relations and bringing to light actors who are currently not seen or do not have the power to design development initiatives or evaluate them. As defined by one interviewee from Africa, systemic change refers to bringing those excluded from conversations on development to the table and engaging them in discussions about their needs and priorities. For a Latin American participant, this attitude differs from welfare policies in developing countries which do not aim to empower vulnerable people in a way they can be inserted or reinserted socially and economically. These unbalanced power relations have led to inequality, especially in countries in the Global South. A Latin American participant, for example, said inequality is a core issue in her region and is related to those unequal power relations that perpetuate the status-quo. This evidence sheds light on the fact that the significant issues the planet is currently experiencing are caused by differences in historically established power relations that have been perpetuated. The concentration of economic and political power is one of the leading causes of inequalities, social injustices, and the planet’s environmental imbalance. Systemic change involves shifting power relations to empower marginalized actors excluded from conversations on development and engaging them in discussions about their needs and priorities. It is in this context of imbalanced economic and political power that most evaluations take place. Evaluations are commissioned by organization leaders, who hold power and define the level of stakeholder engagement.

Systemic change can be transformative or not. A practitioner from Oceania explains that when a systems change is transformative, it should be able to demonstrate a fundamental
difference between how the system behaves or looks before and after the change has happened. She adds that transformative change means significant changes that reshape and are harder to undo. Also, for her, transformative change involves transformative learning, through which people start seeing things differently. She reflected on how changes can be fragile, especially in the public sector. According to her, “When a government starts, there is usually a commitment to an agenda, and people think real progress has been made, then it gets pushed back, or there’s a transformation in a negative direction. Transformations can be harmful, not always positive.”

(PR3-OCE). So, changes in systems do not necessarily lead to significant differences between the way those systems’ function, and the way people see their functioning, before and after a set of interventions. These changes might be a set of incremental or occasional changes which does not make them transformative. Evaluation can help in identifying whether system change was transformative or not by assessing the extent to which a set of initiatives have led to fundamental differences in the way the system behaves or looks before and after the interventions. Evaluation can also assess whether the set of interventions are significant, reshaping, and harder to undo. Evaluation can also determine whether the changes are fragile and subject to being pushed back or transformed in a negative direction. Incremental or occasional changes might not make the changes transformative. Therefore, by evaluating the changes, it is possible to determine whether they have been transformative or not. Finally, evaluation can identify whether transformative learning has occurred, leading to people seeing things differently.

Evaluation plays a critical role in contributing to transformative changes and evaluators are vital actors in this process. This idea was a growing consensus among interviewees from different regions and contexts. One interviewee, who works for a development bank, reflects that beyond measuring progress and outcomes over time, evaluation can be used to identify and
understand the root causes of complex problems, which can inform decision-makers about the most effective solutions. It can also provide feedback to stakeholders, including decision-makers and practitioners, which can be used to adjust strategies and tactics to achieve the desired changes. Another interviewee from Africa points out that the process of evaluation, which involves collecting and analyzing evidence, making conclusions, and justifying decisions, is highly structured. From his perspective, this process is intriguing and has the potential to bring about transformative changes in systems. Even though the process is potentially helpful, changes in systems depend on leadership decisions, which evaluation by itself cannot guarantee. A different interviewee believes the evaluation community is getting more conscious about their roles as advocates for change.

We [ourselves] are activists. So, we can’t distance ourselves imagining that we are just purveyors of evidence and somebody else is responsible for utilization. We need to show that we are adding value. And I am seeing that shift happening. Nobody’s talking about evaluation, just for evaluation’s sake anymore. (AC4-AFR).

A leader of an international VOPE explains with a specific situation. She states evaluations can play a role in improving aid efforts in response to contexts of displacement, lack of resources, and lack of security. Evaluations are in general commissioned by organizations according to pre-determined Terms of Reference. Even though normally there is not much margin to change the scope of work, evaluators can investigate the organizational mechanisms that lead to the use of findings, engage decision-makers in the evaluation process, and make recommendations for effective use. So overall, evaluation can be a powerful tool for driving systems change by providing the information and evidence needed to understand problems,
assess their complexity, identify causes, measure progress, design solutions, and persuade decision-makers to act for transformative change.

Evaluations must consider several elements to contribute to transformative change effectively. The first element is the societal issues the planet is facing. A leader of an international VOPE said evaluation should work towards a more comprehensive approach with the potential for contributing to a more peaceful, just, and sustainable world. From her perspective, evaluation has a role, for example, in improving aid efforts and addressing the needs of individuals suffering or experiencing statelessness. For another interviewee, given the interconnectedness of natural and human systems, evaluations can promote positive change for individuals and the planet. Societal issues such as inequality, discrimination, poverty, and climate change characterize contexts underlying many of the challenges faced by communities and must be considered in all evaluations. They are often the root causes of problems that require change. To contribute to solving these issues, evaluations must understand and acknowledge them. Furthermore, evaluations that consider societal issues can help identify the most effective interventions that address the root causes of problems, rather than just treating symptoms. By evaluating programs and policies through a lens that considers social and cultural factors, evaluations can help ensure that interventions are designed to meet the needs of the communities they serve and are sensitive to local values, norms, and traditions. This approach can promote community ownership and participation, leading to more sustainable and transformative change.

Another element of contribution to transformative change is the institutionalization of national evaluation systems in the countries. A representative of an international VOPE points out that even though there have been several influential evaluations, evaluations of isolated initiatives are less likely to contribute to systemic changes. In her opinion, evaluations that occur
within national systems and interconnected frameworks can be more impactful. If evaluations were conducted with a more systemic view in the countries, they would have a much more significant potential to contribute to transformation because their results would be used to advocate systemically for better policies. In her opinion, in the national public policies context, the office of the presidency is the best place to allocate the evaluation function instead of only in sectorial commissioning units. However, even with a strong argument in favor of systemic evaluations in the countries, there are some tradeoffs:

The question is: do we need all the countries to have their [evaluation] systems? And if they don’t have the systems, cannot they afford to talk about evaluation in a transforming world? Not necessarily, nothing is black or white, but if they don’t, the contribution is going to be less systemic, less tangible, and less relevant, less impactful if the systems are put in place. So that is the idea. (IV1-LAC).

This same interviewee reported that Niger had established a national evaluation policy, is creating a national M&E system, and the government is inviting all actors to participate. So, the idea is structuring evaluation as a function, contributing to public policy design and implementation. Institutionalizing national evaluation systems in countries can contribute to transformative change by promoting deeper and more fundamental shifts in the way governments operate and how they serve their citizens. The first advantage would be to allow a system-wide view of public policies. For example, instead of evaluating just individual programs, national systems could be designed to evaluate entire systems, such as health or education systems. This can lead to systemic change, as evaluations can identify cross-cutting issues that require coordinated efforts to address. The second advantage is that evaluation can produce evidence about what works for whom and in which circumstances. This type of evidence does not result
from one evaluation but from national systems, which can influence national policies. Finally, national evaluation systems can also serve as a platform for testing and scaling innovative policies and programs. By testing and refining these policies, programs, and services, countries can promote transformative change by adopting new and innovative approaches. In summary, institutionalizing national evaluation systems in countries can contribute to transformative change by promoting system-wide change and policy innovation. By encouraging countries to adopt new and more effective ways of delivering services, national evaluation systems can help foster lasting and transformative change.

A last element resides in the advocacy role evaluation can play for change. Several interviewees (e.g., IV1-LAC, IV4-LAC, DB1-NAM, AND PR3-OCE) have emphasized the importance of evaluators taking an active stance in advocating for change by, for example, addressing the needs of marginalized and underrepresented groups. Evaluators can also be political actors who can provoke learning and influence political decisions by providing evidence. Overall, the interviewees said that the evaluation community should focus on advocating for the inclusion of societal goals, such as climate change, gender equality, and the inclusion of marginalized groups in evaluation processes. During an evaluation process, evaluators have the power to propose evaluation questions that can provoke reflections about who is being left behind from the different phases of evaluation processes. Evaluators could also leverage opportunities like conferences, research, journals, and guidelines to affirm their values and principles. Evaluators have power of influence and must recognize their responsibility to not only provide evidence and respond to Requests for Proposals, but to actively work towards promoting positive change and adding value to society.
How Evaluation Should Change to Drive Transformative and Systemic Changes.

Current evaluation practices are not effectively driving transformative and systemic change. From the perspective of some of the interviewees, a reinvention of evaluation practice is necessary to fulfill evaluation potential for contributing to social change and justice. The common sentiment was that "we cannot do business as usual." One of the participants argues that it is becoming clear, especially after the COVID-19 pandemic, that all the systems established to gather evidence around programs and policies are not always helpful and valuable, highlighting the need for evaluation to become more responsive to a rapidly changing environment. So, the traditional way evaluation has been conducted over the last several years is not meeting the needs and expectations of the ones who need to make decisions fast in complex and changing environments. Therefore, to bring about transformative and systemic changes, evaluations must undergo transformation in multiple aspects.

Countries must build up local evaluation ecosystems. Two interviewees described the importance of increasing the representation of individuals from the societies and countries evaluated. From one of the interviewees’ perspective, a baseline issue is the need to build up evaluation ecosystems in regions and countries that are still starting to evaluate their initiatives. Building local evaluation ecosystems involves a range of efforts, including training, technical assistance, research on evaluation, policy development, and strengthening systems that include higher education institutions, private consulting firms, statistical agencies, and government frameworks. The interviewees added that building local evaluation ecosystems will improve evaluation processes’ overall efficacy and ownership. An evaluation ecosystem is a network of interconnected stakeholders, institutions, policies, and practices involved in the evaluation of programs, policies, or initiatives. This ecosystem includes various actors such as evaluators,
stakeholders, policymakers, and institutions engaged in evaluation processes. An evaluation ecosystem also involves the development of necessary infrastructure, resources, and capacities to carry out evaluations effectively. Overall, a robust evaluation ecosystem aims to promote accountability, learning, and improvement in the design, implementation, and evaluation of programs and policies. Building evaluation ecosystems can contribute to making evaluations more transformative in several ways. First, it leads to capacity building of individuals and institutions to carry out better evaluation practices, generate more accurate data, and make more informed decisions. Secondly, it can promote ownership and sustainability of evaluation processes by involving local stakeholders in the evaluation process. This can lead to greater buy-in and commitment to the evaluation process and its outcomes, making it more likely that evaluation recommendations will be implemented. Lastly, building evaluation ecosystems can promote a culture of learning and adaptation, where stakeholders use evaluation findings to improve program or policy design and implementation. This can lead to transformative change by facilitating continuous improvement and innovation, leading to more effective programs and policies over time.

Evaluation practice must change towards better theories of change, context responsiveness, and attention to equity and environmental sustainability issues. During the interview, an evaluator from Oceania discussed the implications of the changes that need to be made to the practice of evaluation. From her perspective, better theories of change, context responsiveness, and attention to equity and environmental sustainability issues must be intrinsic to everything evaluators do.

Evaluation processes need to help evaluators make sense of what’s going on and develop strategies to address the challenges. And this approach needs to come down to all the
ways evaluators do things. It’s not about whether we should look at equity or the environment. Both. Should we look at systems thinking, complexity, or environment? Both. They’re all coming together because that’s the sort of transformation we need. And how do we talk about this stuff in ways that people can understand and engage, and they see the value in it? That’s, to me, the big challenge. That’s what graduate programs should be hoping to do, and we don’t know the answers yet. (PR3-OCE).

Effective theories of change must be intricately linked to the systems and contexts in which the evaluation objects are inserted. One interviewee, a leader of a development bank, exemplified how systemic thinking can be incorporated into evaluation practice:

First, it is necessary to understand a system, for example, the education sector in a country. This system can aim to improve access to education or the quality of education. In both cases, it is essential to understand the root causes, the dynamics of this system, and the intended outcomes and how they interact. Secondly, by understanding the system, you identify some constraints to generate the intended outcomes, which could be more children in schools or better performance of students in schools. And then, you implement different interventions and evaluate them. So, systemic change means that there is sufficient scale, or the right catalytic interventions, that sustainably can change the situation in the system. And sometimes the combination of that, in our case, needs to sustainably be having more children getting into schools or sustainably pushing the boundaries of quality of education and student performance. (DB1-NAM).

Because the first demands were to evaluate programs, evaluation theories were centered on program evaluations for a long time. So, evaluations were historically practiced in silos, independently and individually, using linear theories of change without any connection with
larger systems. According to an interviewee from India, even though evaluating isolated objects is still valid and can lead to process learning and initiatives improvement, this learning process should be extended beyond a project boundary so that more stakeholders can understand the practical challenges and barriers within the system. He adds that this understanding depends on the connection between different programs and different stakeholders and that systems analysis can help operationalize complexity. System analysis looks at complex situations as a series of interactions in multiple dimensions, each of which may be measurable to some extent. Consequently, the way evaluators formulate theories of change must be altered. Linear and one-dimensional models, currently dominant in evaluations, don’t reflect the complexity of the interventions and don’t provide a holistic view of the situation.

Incorporating systemic thinking into evaluations can change the purposes that have guided evaluation in most institutional settings. For some interviewees, evaluations, especially the ones commissioned by investors and donors, are still mostly compliance-driven and output-oriented. Consequently, the data produced are primarily about what has been achieved. For them, if evaluations keep being limited to this type of purpose, it will not make the field relevant or sustainable, and instead will run the risk of the field become archaic and redundant. From the perspective of another interviewee, development agencies and funders need to assess systems change, which would be especially useful when considering tradeoffs and synergies between different areas. Compliance-driven and outputs-oriented evaluations are limited because they do not produce ideas about what needs to be done and what is more likely to work in the future. When evaluators incorporate systems thinking into their evaluations, they shift their perspective from looking backward to looking forward. This change in perspective amplifies the potential benefits of evaluations by increasing their capacity to facilitate transformative change.
Incorporating systems thinking in evaluations will also change the way of formulating evaluation questions and the definition of designs, methods, and approaches. Additionally, there will be a transformation in the process of value attribution. According to one of the interviewees, a seasoned evaluator, the conventional approach of attributing the success or failure of an intervention to a single variable has hindered efforts toward bringing about systemic change. Instead, it may be more productive to focus on collaboration and contribution, where the contributions of multiple stakeholders are acknowledged and appreciated in promoting sustainable change. Hence, incorporating principles of complexity and systems thinking in evaluations can modify the evaluators’ perspective on impacts and causal relationships. As incorporating systems thinking can lead evaluators to consider a wider range of factors that may impact the system, questions will be broader. Questions that capture the complexity of the system being evaluated will require different tools and techniques. The process of value attribution will consider the various ways in which value is created and distributed within the system. Overall, incorporating systems thinking in evaluations can lead to more nuanced and comprehensive evaluations that better reflect the complexity of the systems being evaluated.

Another crucial change necessary in evaluation practice is the incorporation of context. Incorporating context in evaluations means considering the extent to which various circumstances impact how evaluations are conducted. One of the interviewees shared her experience as a peer reviewer in the evaluation process and the challenges she faced. She mentioned an evaluation of a regional strategic plan commissioned before the outbreak of COVID-19 and implemented during the pandemic. In this case, the interviewee believes that there was an ethical obligation to consider the impact of the pandemic on the evaluation and to adjust recommendations accordingly, even if it entailed acknowledging the altered context and
incorporating relevant information about the pandemic into the evaluation. The interviewee also emphasized the possibility of evaluations being impacted by sudden and substantial changes in context, such as wars, and the ethical responsibility to consider these changes and adapt accordingly in the evaluation process. Incorporating context into evaluations involves several aspects. One is to consider the characteristics, history, habits, and values of the people or organizations involved. Another is to consider any changes in the evaluation context that may have occurred during the design and implementation of some initiative, as well as any changes that may arise during the evaluation process. Considering context in both aspects can influence the level of engagement, the formulation of evaluation questions, and the selection of data collection methods and tools, as well as the way in which value is attributed. Evaluation must move from predetermined models and the concept of a golden standard to adopting flexibility, adaptation, and resilience.

Evaluations must include insights and voices of those who haven’t been heard. Another significant change in the current evaluation practice highlighted by participants is the need for a more stakeholder-centered approach. A stakeholder-centered approach entails engaging with relevant stakeholders and ensuring that all parties have a voice and influence in the evaluation process, making it more inclusive and impartial. One of the interviewees provided an example of how accounting for the context of the COVID-19 pandemic was essential when engaging stakeholders.

A lot of research was done on the different implications and consequences of COVID. But a great deal of the research left out minorities, African Americans, and women to some extent. That’s still because there are a lot of well-organized academic women who sort of make sure they get a seat at the table. But that’s just an example where many
people thought they were doing rigorous research on COVID, but just didn’t think that they were leaving out the effects on the Hispanic or the African American population or other vulnerable groups. (PR2-NAM).

By including stakeholders in the evaluation process, valuable perspectives and insights can be obtained, leading to more comprehensive and accurate evaluations. Additionally, stakeholder involvement can enhance their understanding of the initiative’s outcomes and evaluation process, fostering greater trust and support for using evaluation findings. Furthermore, stakeholder engagement can result in effective changes and improvements, as they are more likely to support initiatives they have helped to shape and align with their needs and interests. So, by including insights and voices of those who haven’t been heard, evaluations are more likely to address inequalities.

**Themes or Topics Likely to Gain Importance for Evaluation Practice in the Coming Years.** As evaluation practice incorporates changes, some themes or topics are likely to gain importance in the coming years. The participants were queried on their perception of these themes or topics. Most respondents indicated that sustainable development and climate change are likely to become critical foci of evaluations due to their urgent nature. Related to these topics, there is a growing recognition of the essential role that indigenous knowledge systems can play in finding solutions to address climate change as for those systems preservation of the environment and promotion of sustainable living are core values. However, is not common to see in practice the incorporation of indigenous philosophy in evaluations, making indigenous communities one of the groups that are excluded from this kind of process. The advocacy for the inclusion of indigenous knowledge systems in evaluation is part of a more significant movement towards decolonizing evaluation, as the incorporation of indigenous knowledge systems is not
commonly seen in Western evaluation practices. A group of researchers led by African scholars is investigating the integration of indigenous philosophy into the evaluation practice as a means of decolonizing it. One interviewee from Africa stated that there is a close relationship between indigenous knowledge systems, complexity, and systems thinking in African indigenous culture. According to her, recent writings reveal that indigenous knowledge systems are and have always been complex, in contrast to Western knowledge system, which has always been linear. Indigenous knowledge systems have always been centered around systems thinking, as they believe in the interconnectedness of things and that people are not isolated from the environment. However, this was not previously referred to as systems thinking. The evaluation community is starting to recognize this and is exploring new methodologies for comprehending complexity in evaluation. Therefore, sustainable development and climate change are critical issues of the planet and international agreements and agendas have stressed the need to reduce carbon emissions, stop global warming and regenerate the environment to guarantee natural resources for our future generations. The evaluation community must become aware of and understand the root causes of these issues and the factors that can provoke changes, as well as provide evidence on what works in which circumstances and contexts.

Another issue brought by one of the respondents, who has extensive experience working for UN agencies, is conflicts and their economic and social impacts at both the local and global levels. She believes that the future of evaluation is closely tied to peace and the global consequences of crisis affecting people’s lives. From her perspective, evaluations conducted in the context of crisis and violence must be connected to ethical inquiry, ensuring that beyond meeting ethical protocols and requirements, they are sensitive to the tensions and dilemmas faced by stakeholders in real-world work. These tensions and dilemmas are generally related to
fundamental power relations and ethical issues in life. In this sense, the evaluation community must reflect on the extent to which evaluations in areas of conflict consider the new marginalized populations, such as refugees, immigrants, and orphans, and the social impacts of those conflicts, such as poverty, citizenship status, and hunger as a context in their evaluations. By shedding light on the negative impacts of conflicts on people’s lives, evaluations can support peace negotiations and decisions about funding, punishing, or discontinuing conflicts.

To have a just and fair planet, a systemic change must happen that prioritizes fighting all types of injustices and inequalities. Social justice and human rights (gender equality, violence against women and children, equity, poverty, race, etc.) were some of the themes brought by the respondents as likely to gain relevance to the evaluation field, as they may represent the core issues for the less developed regions in the southern continents. Participants believe evaluation can help address inequalities, community discrimination, and exclusion. By making explicit how factors that cause injustice, inequality, and the violation of human rights are systemically present in organizations and initiatives, evaluations can prevent the reproduction of power imbalance and the perpetuation of the status quo.

One of the interviewees also mentioned the workforce as an essential topic in the future. According to him, the composition and demands of the workforce in the United States were directly impacted by the COVID-19 pandemic. After most of the countries decided to lock private and public businesses and services down, isolating the population in their own houses, all economic activity had to immediately become remote, dramatically changing the relationship between employees and employers. People had to start working and studying from home, taking advantage of technological advances. This same interviewee also points out that the pandemic highlighted significant inequities and brought issues such as housing, homelessness, and the
The use of technology has become an important topic in the field of evaluation. One of the interviewees, a coordinator of an evaluation education program, points out that the COVID-19 pandemic has had a significant impact on the use of remote technologies, such as video-conference applications, which have become ubiquitous tools for communication. Remote technologies have enabled evaluators to save money and time by eliminating the need to travel for interviews and meetings. The power of cooperation has also increased with remote technologies, as evaluators can collaborate and exchange documents and knowledge easily. This interviewee also emphasizes the importance of lifelong learning and how online courses have made it easier for evaluators to continue learning and improving their skills without having to attend classes on campus. The use of online courses has enabled evaluators to focus on their specific needs and level of expertise, making progress at their own pace. Furthermore, the
academic environment has become more international due to technology, making it easier for evaluators to exchange information and collaborate with people from all over the world. The interviewee also mentions that technology has made it possible to teach techniques and methods more easily online. Language is no longer a barrier as courses can be offered with subtitles or in different languages. Thus, technology has enabled evaluators to extend a broad effort to everyone, ensuring that information is accessible to all. These benefits are crucial for evaluators who need to keep up with the constantly changing landscape of their field.

The salience of these themes to evaluation practice will vary depending on the institutional context and cultural norms. Some interviewees point out that some organizations are more inclined towards incorporating a focus on equity and sustainability into their evaluations. However, this may not be a widely accepted practice, particularly in large evaluation systems within institutions such as the European Commission, multilateral development banks, and national governments. They added that despite scholarly literature highlighting the importance of these themes in evaluation, there remains a discrepancy between research and the actual implementation of these principles in evaluations conducted within these larger systems. So, the level of acceptance and implementation of these principles may differ depending on the organization’s culture and norms, which can be influenced by various factors such as political and economic interests, history, and values. Therefore, institutional contexts and cultural norms may affect the extent to which equity and sustainability are incorporated into evaluation practice.

**Approaches Likely to Gain Importance for Evaluation Practice in the Coming Years.** During the interviews, seven approaches were mentioned as likely to gain importance for evaluation in the near future. First, most participants emphasized the significance of incorporating complexity-responsive and systems-oriented approaches in evaluations. From one
of the interviewee’s perspective, the systems-oriented approach involves mapping the system initially, analyzing the relationship between crucial outcomes and causal factors, and examining the impact of different types of interventions on these causal factors, given certain conditions. Another participant, a practitioner from North America, said that to incorporate systems thinking in evaluations, evaluators must comprehensively understand the economic and political structural transformations shaping society and the growing interconnectivity among various sectors. She provided an example of the perspective of economic justice related to domestic violence and domestic violence prevention. Different social determinants are connected to the perpetration of domestic violence, such as criminal justice systems, living wages, and access to health care. She points out that many of those problems are clustered in certain communities because of the inequities. For her, evaluation can support the transformation of those systems to become more equitable. “If we could improve these conditions, the likelihood of the perpetration happening would be greatly reduced. And take the spotlight from the individuals and put it on the systems. So yeah, that is transformative. Right?” (PR4-NAM). Another systemic aspect mentioned by this same interviewee is assessing how organizations work together, which could add value to complex evaluations. “It is a value addition to them working together. So, it’s not just traditional questions, there are some complicated questions, and it requires a lot more thinking about complex ideas and how to break it down to something manageable.” (PR1-NAM). Evaluation can do much more if processes are collaborative and inclusive to make sense of data. So, the incorporation of complexity responsiveness and systems thinking in evaluation is important because they can provide a holistic view of the parts of a system and the interconnections between these components and outcomes. Also, the understanding of the context in which these systems operate and how it influences these interconnections. In a world facing rapid changes in
society, systems thinking enables evaluators to better understand transformational change and its implementation.

Second, participants cited collaborative approaches as essential. Empowering, participatory, or deliberative approaches ensure that different groups have a voice and learn with the evaluation process. One interviewee from Latin America points out that participatory approaches must go beyond queries based on pre-formulated questions. From her perspective, interviews and surveys can still be used. However, participatory methods imply better communication with participants aiming to guarantee they understand what is being done and why. According to her, people engaged in evaluation processes can give the program leaders their perspectives on what is needed to solve the problems. She added that participatory approaches could help build an understanding and acceptance of the need for accountability and transparency, especially in her region. Another Latin American interviewee sees exceptionally high value in the empowerment evaluation approach because it lets people be accountable for the success of what they do. So, participatory approaches can help engaged groups understand the relevance of actions and ideas, what the evidence is suggesting, and what the choices are, and make them reflect and share thoughts about what is important to them, influencing decision-makers to change the way they see solutions. These kinds of approaches are more likely to provide meaningful, sound, and appropriate findings that could drive social change.

Third, the utilization of the developmental evaluation approach has the potential to instigate substantial and pervasive transformations within a system. From the perspective of a Latin American interviewee, the developmental evaluation approach fosters cyclical and adaptable procedures, nurtures experimentation and acquisition of knowledge, encourages cooperation and involvement of stakeholders, and continually furnishes feedback to guide
decision-making processes. This approach emphasizes adaptability and responsiveness to alterable circumstances. Substantial and pervasive transformations within a system depend on a deep understanding of how the parts of the systems influence outcomes, and the acknowledgment that context continuously changes, requiring adaptations, experimentation, and action along with the implementation. It is a cyclical process of acting, collecting evidence, and changing directions that occurs during the implementation of innovative initiatives. Learning and improvements happen along the way. Adopting a developmental evaluation approach enables organizations to effectively tackle complicated and rapidly evolving systemic challenges and make the necessary modifications to attain their intended results.

Fourthly, it was noted that contextually responsive approaches hold the highest potential for transformative evaluations. The evaluation paradigm must shift towards an approach not imposed on individuals but conducted by them. According to a scholar from Africa, evaluations in developing nations should be contextually responsive and leverage indigenous philosophies by incorporating methods, techniques, and strategies pertinent to the local context. She states that evaluations must generate empirical data that address questions relevant to the specific contexts. The results of the interviews indicated that certain regions are exploring methods that are more attuned to their regional contexts, such as the "Made in Africa Evaluation" approach. Nevertheless, some interviewees still have some reservations regarding this concept’s exact nature and potential. The Made in Africa evaluation approach is an example of the positioning of the evaluative community in a region that argues that, for evaluations to be transformative, they need to consider the contexts where evaluation objects operate. A contextually responsive evaluation brings an accurate view of the causes of problems and the most appropriate solutions
according to the local reality. Culturally responsive evaluations carry the local perspective of needs, expectations, and values and are therefore more likely to bring about change.

Equity-focused evaluation was the fifth type of approach highlighted by the interviewees. Evaluation must identify and address disparities in outcomes and access to opportunities. One interviewee, a practitioner who worked for many years for a development bank, thinks evaluators are usually unaware of the limitations of program implementation. For example, a program might be designed and work well for the better of families on average. Still, there might be a significant social exclusion from the program in different contexts, such as low-income families, ethnic minorities, or landless individuals. However, according to what he observes, evaluations cannot capture that because they mostly rely on project records as sources of evidence, which only have data on the people using or benefiting from the projects. By relying just on project records that are not necessarily produced over data collected in the field, evaluations are potentially using biased evidence. This bias prevents an accurate understanding of how projects work in the real world, leading to biased decisions. So, evaluations must include aspects related to inequalities and questions that can examine who is being left out of the initiatives being evaluated. Consequently, the data collection methods and the analyses must be able to reveal and communicate such disparities.

A sixth type of approach mentioned as crucial was Utilization-Focused Evaluation (UFE). Evaluation results must be delivered in a timely manner, as they can be helpful for decision-makers, who represent the users. Some interviewees reported that it has been consistently challenging to deliver timely evaluation results. The COVID-19 pandemic brought the issue of uncertainty and the need to be flexible. New approaches, such as utilization-focused evaluation, have been tested and used to fulfill these needs and deal with these contexts. A
seasoned practitioner from Australia provided an example of how time became an issue in one evaluation.

There was an evaluation before the pandemic about social services in Australia. Then, a massive lockdown happened, and a massive change in how social services were delivered. So, services related to maternal and child health, always developed on-site, had to be immediately turned to virtual. Then, the evaluation had a six-week turnaround. So, the change was not just about collecting data quickly but doing the whole turnaround and engaging people in the questions in another manner, helping them to get good and actionable questions and understand the limitations. The team pointed out that they could do it quickly but would not be able to show the long-term effects. So, they had to manage uncertainty. (PR3-OCE).

In contexts of uncertainty and rapid change, it is vital to understand that decision-makers cannot always wait for evaluation results. Evaluators must be creative and use evaluation designs that are able to collect the best evidence possible with the resources available, including time. In these situations, evaluators must be able to realize how evaluation can provide strategic data of what matters to lead to transformative change.

The seventh and final type of approach deemed essential was those that allow for the evaluation of environmental programs or approaches that integrate environmental aspects in all evaluations, such as the ‘Footprint Evaluation’.

Methods Likely to Gain Importance for Evaluation Practice in the Coming Years. The combination of quantitative and qualitative methods in evaluations can provide a more comprehensive and nuanced understanding of the program’s outcomes. For one of the interviewees from North America, the evaluation community has been heavily influenced by the
research community and has held the belief that only quantitative methods could provide valid evidence. However, he thinks this exclusive reliance on quantitative methods has been widely acknowledged to have several limitations. For instance, it can suffer from the unavailability of data, errors, and biases, thereby precluding meaningful subjective experiences or perspectives. He adds that quantitative data may not equip teams to effectively evaluate complex or multi-faceted programs nor accurately capture a program’s long-term impact or unintended consequences. While some people maintain that traditional and rigorous methods are the only acceptable means of data collection, the need to evaluate and monitor progress towards complex objects, such as the Sustainable Development Goals of the 2030 Agenda, necessitates a reconsideration of these views. According to him, the use of big data in evaluation has gained popularity due to its potential to expedite evidence collection and facilitate quicker decision-making. However, it is essential to acknowledge the potential biases in big data and consider the limitations and strengths of these data collection methods and cultural biases present in the data, to ensure an accurate representation of reality. He provides an example:

Specialists who use data science always give examples of satellite images used to measure some phenomena introducing the concept of ground truthing. They believe it is possible, for example, to measure poverty by looking at the quality of construction through images. When construction is poor, this means it must be a low-income family. And frequently, these analyses are made from their offices. However, it is fundamental to check on the ground whether the things being interpreted from the images are correct. And you constantly find that many families who live in impoverished housing are not necessarily poor. There are many reasons why they don’t want to invest too much in the house. (PR2-NAM)
He continues by arguing that remote data collection has demonstrated efficacy in certain situations, especially during the COVID-19 pandemic when conducting in-person fieldwork has been challenging. Nevertheless, it is imperative to acknowledge the limitations of this method, particularly concerning comprehending complex issues such as domestic violence or community crime, which necessitate direct observation and engagement. Additionally, there is a danger that decision-makers not directly involved in fieldwork may not possess a comprehensive understanding of the experiences and needs of the communities they aim to serve. So, although there is a belief that quantitative methods are the only ones capable of providing reliable evaluation results, several situations such as those presented by the interviewee show that these are not always the most adequate methods to provide valid evidence. Another piece of evidence presented was that in research regarding the effects of COVID-19, there has been a bias toward overlooking the effects on minority and marginalized populations, leading to an incomplete or skewed comprehension of the pandemic’s ramifications. Hence, it is imperative to consciously endeavor towards inclusiveness in research endeavors, to accurately capture and address the experiences and consequences experienced by all members of society.

Given the limitations of quantitative methods that were pointed out, the interviewee thinks donors and other stakeholders need to prioritize spending time in the field to fully understand the realities on the ground and make informed decisions. Qualitative methods are closely tied to the principle of ground truthing, which involves verifying the validity of data interpretation and comprehending how programs operate.

The interviewee cautions against the tendency to rely solely on satellite imagery, as it raises questions about the accuracy of interpretations from a distant vantage point. He provided an example. He formerly served as a specialist in housing programs at a development bank. It
was deemed noteworthy to encourage employees from the housing agencies and funders to inhabit one of the housing projects they were overseeing. Before this initiative, no one from the bank had resided in the houses they were financing globally in order to better understand the recipients of their efforts. Some personnel agreed to live in a slum temporarily, leading to a profound and life-altering experience. They were astounded by the discovery that the inhabitants of the slum were not lazy but possessed a diverse array of abilities, including construction, woodworking, and culinary skills. In addition, the formidable nature of the women in the community, who were responsible for maintaining their family’s cohesiveness and actively protesting against the government’s plans to demolish their homes, left a lasting impression on them. The experience challenged their prejudices and instilled a deep sense of respect for the resourcefulness and resilience of the residents. The initial intention of the initiative was to assist what was perceived to be a marginalized population. Still, it ultimately proved to be an eye-opening and transformative experience for the participating employees. So, the utilization of qualitative methods in evaluation is the only way to capture the richness and complexity of the realities on the ground. They allow evaluators to understand the lived experiences of program participants, by capturing their perspectives, attitudes, beliefs, and values. Consequently, they can provide insights into how programs can better meet the needs of participants and stakeholders. Lastly, qualitative methods can increase engagement by providing opportunities for stakeholders to participate in the evaluation process and share their perspectives. This can increase buy-in and support for the program and the evaluation. Overall, the use of qualitative methods in program evaluations can provide a more comprehensive understanding of program processes and outcomes and can facilitate program improvement and stakeholder engagement.

One of the specialists interviewed has tried to improve the quality and use of data in evaluations,
in an organization where most of the people come from an economics background. This organization has focused explicitly on incorporating qualitative data and mixed methods, as well as process evaluation, which aims to understand better how programs are implemented. The interviewee notes that many impact evaluations, particularly those using randomized controlled trial designs, do not consider the implementation of projects and, instead, focus solely on the project design and outcomes. This limitation can lead to the conclusion that a project was unsuccessful due to a flawed design when there may have been implementation issues.

Based on these limitations, most interviewees advocate using a combination of quantitative and qualitative methods, or mixed methods, to provide a more comprehensive and nuanced understanding of the program’s outcomes. So, even though people have a growing interest in the data generation side of the evaluation business, there is a need to develop the interpersonal side for continuous data validity. And this interpersonal side can be brought about by the engagement of different groups in the evaluation process, which must have the opportunity to give their interpretation of the data collected. Moreover, as we become increasingly dependent on big data, it is essential to understand and educate evaluators and other users on the nature of big data—both its strengths and some severe limitations. Therefore, mixed methods were the most recommended by the interviewees. And for one of them, it also involves being a mix of participatory and non-participatory, or being a mix between empowering and non-empowering, as participatory can be very instrumental and have the purpose of just getting information.

**How trends are becoming evident.** Evaluation for transformative change is a new concept and people are at different levels of understanding of how evaluation can contribute to systemic change in different parts of the world. This became more evident by comparing
perspectives between interviewees from the Global North and from the Global South. One interviewee from India perceives some concrete advances in his local context through the growing attention to the learning of programs implementation and to the connection of different programs and stakeholders. But he thinks the change will only come when people practice evaluation more meaningfully. Another interviewee from the U.S. reported that she has seen a lot of transformation in U.S. federal programming in evaluations. She points out that, in the past, using terms such as "evidence base" was not permitted for evaluators. However, over time, the evaluator community has come to acknowledge that, for example, the social determinants of health are inextricably linked to social justice. However, she recognizes that the U.S. needs to move forward on many societal issues. For everybody to be able to live their healthiest lives, the U.S. must think about the transformation of systems and systemic injustices, racism, and all those inequities that have become much more explicit because of the visibility and vocal willingness of society. She thinks evaluations should prioritize sustainability in evaluating programs and organizations and the impact of human actions on the planet for future generations. She has observed a transformation in her work in the Centers for Disease Control and Prevention (CDC) over the past 12 years, including increased diversity among science and program officers. This shift towards a more diverse workforce is expected to result in more diverse decision-making and problem-solving within the organization. So, changes in evaluation practice happen in different ways for different contexts and depend on several factors like level of development and societal values. All the characteristics described by the group of interviewees as important for evaluation practice for transformative change can be seen as guiding principles and probably won’t happen simultaneously in all the contexts.
The integration of critical themes into evaluative practices, which can potentially drive change, is not being evenly adopted by organizations. One interviewee from an African learning center stated that some donors have prioritized areas already identified as crucial, such as the four thematic areas defined by the Global Evaluation Initiative: climate change, fragility conflict and violence, youth, and gender. According to her, even though implementing partners still don’t prioritize those four areas, the fact that they were chosen by the donors and those driving the commissioning of evaluations can potentially drive change. Another interviewee from an international funder noted that integrating sustainability and equity considerations into evaluative practices is limited by prevailing cognitive frameworks and established protocols in organizations like multilateral banks. As a result, this interviewee perceives limited progress in addressing these concerns in his current and previous workplaces. Another interviewee, a leader of an international VOPE, reported that she is not seeing from the demand side a focus on marginalized people in evaluations of programs related to violence and conflict. The slogan “no one left behind,” associated with Agenda 2030, is always present, but it is a broad term. She thinks evaluations of programs related to violence and conflict should consider the orphans, the displaced people, and those who have become refugees or do not even have that status, making evaluations even more complex. However, she observes the migrants are only mentioned in general terms. Another interviewee, who works for a European Think Tank on environment and development, reported that in his evaluations, the sustainability issue remains in the background, even though more questions addressing this topic are emerging.

Commissioners request you to look at the interaction and the linkages between the different programs, how coherent the response is, whether there has been a systemic change. But sometimes, I feel like they are formulating questions to produce findings for
those implementing policies or problems. Even though in practice, there has been a lot of
discussions on climate change and sustainability among the European evaluation
community, especially at the European Evaluation Society conferences, which is a
preannounce of change. (RI1-EUR)

Regarding incorporating complexity, one of the interviewees thinks those discussions are
still limited to the theoretical level, and it has not been very well addressed in many evaluations.
From the demand side, commissioners have asked for more evaluations considering complexity,
but there are still many traditional requests for projects and program evaluations. In the Requests
for Proposals demanding more complex evaluations, terms like transformation and systemic
analysis are present, but expectations regarding timeframe, for example, are unrealistic. These
unrealistic expectations show a mismatch between incorporating these terms and thinking about
the implications for evaluation design and implementation. On the supply side, one of the
interviewees stated that most evaluators do not use this concept a great deal. For him, the
evaluation toolkit doesn’t have many analytical approaches for incorporating complexity.

Regarding systems thinking, some interviewees highlighted the significance of Agenda 2030 and
its related monitoring and evaluation processes in promoting this type of thinking because the 17
Sustainable Development Goals are interlinked. Because this is a severe issue in Latin America,
one of the interviewees uses as an example the gender mainstreaming and evaluation approach
that has been adopted among some groups of consultants. According to her, “it’s not just a class.
It’s not something additional or optional. On the contrary, it should be a quality criterion because
one cannot address systemic change without addressing unequal gender relations and everything
related.” (IV4-LAC). Another interviewee from an international VOPE reinforces this position
and adds that, in her perspective, in addition to being gender sensitive, the equity dimension should be cross-cutting.

The interviewees also reported limited adoption of innovative approaches in the field. One practitioner from Canada pointed out that, from the supply side, there is a lack of systematization in the application of new approaches and difficulty in pricing them in consulting proposals. Additionally, new approaches may not align with procurement processes, often involving mechanical repetition of templates. Adopting innovative approaches requires proposals to be formulated differently and may result in increased costs due to the time and effort required, for example, to engage stakeholders. In terms of new data collection methods, the interviewee who approached this topic said he still does not see many evaluations using big data. However, he observes some use of big data for program design and operational support. Another interviewee reported he observes a greater use of geospatial data, which can be helpful in sustainable development monitoring, even in countries that are not generating this kind of evidence.

So, there were different perceptions among interviewees about the potential to integrate crucial themes, approaches, and methods in evaluation practices. Even though some organizations are adopting initiatives that can stimulate this integration, this is still not an even strategy, and the effectiveness of its incorporation depends on the implementation agencies. International organizations, as providers of funds and technical assistance to several countries, can influence that evaluation practices incorporate the changes to become more connected to a world in transformation.

The evaluation community is increasingly advocating for the utilization of evaluations to drive change. Interviewees mentioned that some groups of evaluators have acted as activists to
demonstrate how evaluation can bring value. According to an African interviewee, evaluations are no longer viewed as valuable purely for their own sake. One of the interviewees mentioned that evaluation used to be an elective course, but now it is part of core curricula. She observes that evaluative concepts and terminology are used even at lower education levels. She reports that her daughters were enrolled in a high school that strongly emphasized experiential learning and provided opportunities for advanced research. She expressed surprise at the level of evaluative instruction provided at the high school, comparing it favorably to her graduate school education. The evaluation community is represented by international, regional, and national VOPEs, present in all continents and in various research and teaching organizations. Through conferences, dialogues, and publications, these organizations engage the various stakeholders that form evaluation ecosystems in the discussion about the role of evaluation to drive change. This advocacy role can provoke reflections and influence the revision of evaluation practice and education.

Finally, the interviewees shared some reflections about how the new trends affect their professional practice. For example, one of the interviewees, a practitioner, compared the way she used to do evaluations 15-20 years ago and the way she is currently doing them. She reports that the most significant difference is that she and her team are currently trying to get better conceptual clarity on specific terms and ideas and paying more attention to how to translate them into practice. Another interviewee from a Development Bank reported that he is involved in the discussion on translating the use of new data and new data analytics, including artificial intelligence, machine learning, deep learning, etc., into evaluative practice. Finally, one European interviewee said that evaluations with which he has been involved still have a high
focus on results. He feels sometimes commissioners want to have auditors on results other than evaluators, which is another job.

**Challenges that Lay ahead for Evaluation as a Field in an Era of Transformational Change.** Several challenges lay ahead for evaluation given its role to contribute to transformative change. One challenge pointed out by the interviewees is to establish evaluation as a transdisciplinary field, which requires evaluation professionals to have cross-disciplinary skills and knowledge. The interviewees highlighted that beyond core evaluation knowledge, evaluators need to be aware of the latest developments in research to produce high-quality evaluations. This awareness requires evaluators to be closely connected to other communities of professionals and to continuously upgrade their skills to keep pace with a rapidly evolving world. Establishing evaluation as a transdisciplinary field is a challenge because it requires breaking down traditional disciplinary boundaries and creating new forms of collaboration and knowledge production that can effectively address complex societal issues. This is important because the contribution of evaluation to transformative change depends on providing evidence-based insights that can inform decision-making processes across a range of sectors and disciplines. By doing so, evaluation can become a more effective tool for promoting transformative change that is grounded in a deep understanding of the social and ecological systems in which we live.

Another challenge discussed is related to the need to address equity and climate change in evaluations, which requires the expansion of evaluation objects beyond project boundaries and the adoption of new tools and approaches. The interviewees identified two aspects of this challenge. First, they point out that it requires that the funders and commissioners, representing the demand side of evaluations, reflect on the importance of these issues in their evaluation policies and in their Requests for Proposals. One interviewee from a development bank in the
U.S. observes the supply side of evaluation may face pushback or resistance from clients or funding agencies who do not want evaluations to have a particular focus on equity or social justice. In his perception, there are limits to how far evaluators can deviate from the terms of reference established by clients or funding agencies, making it challenging to bring a specific perspective or set of values to their work. So, balancing the need to be pragmatic with maintaining one’s values is challenging in this work. The importance of using more inclusive approaches and considering marginalized groups in evaluations was emphasized by another participant from the international aid arena. So, evaluation can be instrumental in helping decision-makers understand the reasons why specific individuals or groups do not use services provided to low-income communities and why initiatives can be contributing to global warming. In summary, addressing equity and climate change in evaluations is important for achieving transformative change, but it is a challenge because it requires a holistic approach that goes beyond traditional evaluation methods.

Related to the previous challenge is the incorporation of systems thinking in evaluation. Two evaluation practitioners expressed concerns about the focus and approaches that have guided most evaluations. The first, a leader from a European think tank on development, highlights the limitations of the evaluation criteria established by the Organisation for Economic Co-operation and Development (OECD), widely adopted by countries and commissioners as a reference for evaluation practice. He argues that these criteria lead to mechanical and overly structured evaluations, preventing the emergence of relevant questions. Instead, he suggests an approach that incorporates a comprehensive examination of different dimensions, including the interconnectedness of power dynamics, trade-offs, and synergies between environmental and social issues.
I see too much talking about system thinking in evaluation and too little talking about the real issue, which is changing the questions. The real issue today is that evaluation has focused on the wrong questions for the last 30 years, if not more since the beginning. Because the focus of evaluation has been, to some degree, on what works or what have we learned, right? But the real challenge is what will work. Sustainability is always underscored in all evaluations because we are obsessed with results, always looking behind. (RI1-EUR).

However, according to another interviewee from a development bank, the incorporation of systems thinking in evaluations requires proper implementation and faces several challenges. Among these challenges, he cited a limited understanding of the systemic interactions required to solve problems, different perspectives on transformation and justice among different groups, and a lack of coordination among different evaluation units. He added that development banks and multilateral agencies are also limited in their focus on projects and programs instead of considering a broader systemic view.

I would like more space to do a more systems-oriented analysis. From my understanding, a systems approach involves mapping the system first and seeing how important outcomes are related to causal factors. Then, see how different types of interventions may or may not influence these outcomes under certain conditions. And I think we do not have the space, incentives, resources, and even methods and capacities to do this properly. But people still see it more as interesting research, instead of a core of evaluative practice. (DB1-NAM).

The COVID-19 pandemic reinforces the need to adopt a systemic approach to problem-solving due to its far-reaching impact across various systems. However, integrating systemic
thinking into evaluations is challenging in countries with underdeveloped national monitoring and evaluation systems. Mitigating these challenges necessitates augmented coordination, capacity building, and the exploration of inventive evaluation methodologies.

Another interviewee, a practitioner from LAC, expresses concern about the growing focus on impact evaluations in the field. She notes that the frequent focus on evaluating impacts after only one year of intervention is insufficient to understand systemic change. The interviewee defends monitoring, evaluation, and learning (MEL) as a more comprehensive approach adopted by many aid agencies worldwide that emphasize the importance of understanding the evaluation process. Both interviewees stress the need to re-evaluate the current evaluation methodologies, which are often centered on interventions and institutions and are biased towards underestimating the complexity of the world in which they are embedded. A more holistic approach, focusing on systemic change, is required to evaluate the impact of interventions and programs effectively.

Building up evaluation ecosystems, cited before as one of the changes needed in evaluation practices, was identified as a challenge, as it implies articulating their components. An African scholar notes that universities ignore evaluation as a discipline. She points out that if evaluation is a discipline, the evaluation community has yet to agree on what this discipline constitutes. In her perspective, advanced knowledge in evaluation is offered in specializations under specific disciplines like management or public health or in graduate programs in monitoring and evaluation. However, she thinks very few universities offer this last option. Evaluation ecosystems consist of organizations, networks, and professionals involved in evaluation, as well as policies, resources, and culture that shape the evaluation practice. Building up evaluation ecosystems is a challenge for evaluation because it requires articulation and
collaboration among stakeholders, such as government agencies, civil society organizations, academia, and communities to create an enabling environment for evaluation. It also involves putting in place policies and initiatives to support professionalization such as standards, guidelines, competencies, and capacity-building programs. For example, because academia is essential in producing research and offering education, it has an essential role in showing how evaluation can be relevant in the current crisis of our time. By building up evaluation ecosystems, evaluation can become a more effective tool for promoting transformative change and contributing to sustainable development.

Making evaluative evidence relevant, sound, and integrated into public policy processes, which requires well-established and well-designed evaluation functions, was also pointed out as crucial by the interviews. From their perspective, the evaluation community must establish policy dialogues with governments and parliamentarians to raise awareness about how evaluations can contribute to significant societal issues, such as climate change and equity.

This is a very important challenge that we should not underestimate because there are still many policy environments where evaluation is marginalized, and there are still many others where evaluation is coopted in such a way that it becomes harmless. So, it’s ritualized, big machinery of evaluation, but they’re not making a difference in terms of influencing processes of accountability and learning in organizations. There are many positive examples, and we should learn from the positive ones and try to push forward on making our evaluative evidence as useful as possible (DB1-NAM).

One interviewee from Germany reported that in her country, new governmental protection rules against the COVID-19 pandemic established in 2021 were based on the results of an evaluation. She considers this a huge step for evaluation to become influential in politics.
This is a challenge because evaluation is still not seen as a tool that can provide evidence-based insights to inform decision-making processes and contribute to accountability and learning in organizations, especially in the public sector. In such contexts, evaluation may become a ritualized and bureaucratic process that does not make a difference in terms of influencing policy outcomes. By establishing policy dialogues with governments and parliamentarians, the evaluation community can raise awareness about the value of evaluation and its potential to contribute to significant societal issues such as climate change and equity.

The political aspects of evaluating public policies and programs also present various challenges for decision-making based on evidence and evaluation results. One interviewee from the Global South noted that political decisions in democratic systems are shaped not only by evidence but also by political agendas. Evidence’s low influence on public policies can be attributed, in part, to the absence of effective communication tools and authority for those responsible for conducting and conveying evaluation results, which hinders their ability to reach decision-makers. In addition, decision-makers often view evaluations as costly and time-consuming, leading to a short-term mentality in the evaluation process. One interviewee from Latin America exemplified this by recounting an experience with a minister who only had time to read the executive summary of an evaluation report, causing a limited view of the analysis provided. A practitioner from Canada emphasized the need to balance expectations and available resources for evaluations and to adopt new approaches, such as developmental evaluation, to address issues of superficial coverage and poor resource allocation. Behind the evidence provided resides the fact that public policies often face complex environments where stakeholders may have competing interests and may have different perspectives on what constitutes success or failure. Because of this, they might want to influence the evaluation
processes to meet their interests or protect themselves from negative findings. So, political factors influence how evaluations are conducted, interpreted, and used. It is challenging for evaluators to navigate these complex dynamics and political interference and at the same time preserve their independence. Biases caused by this kind of interference can limit the potential of evaluation to promote transformative change by undermining the credibility and legitimacy of the evaluation process.

Another challenge discussed by some interviewees was the influence of institutionalization models in the countries. One interviewee from Africa observed that the national treasuries and finance ministries play a crucial role in shaping the institutionalization of evaluations, often leading to an emphasis on economic foci. This means that evaluations are focused on economic outcomes rather than social or environmental impacts. Likewise, the influence on evaluation approaches and methods was also discussed. A practitioner from the international aid arena mentioned the shift from participatory methods in the 1990s to remote data collection methods in recent times. This shift implies that evaluations may not always capture the voices and perspectives of the communities being evaluated, leading to incomplete or biased evaluations. As mentioned before by one interviewee, evaluations often rely on monitoring reports produced by implementing agencies, which have a positive bias and do not collect data on those who do not use a project. The interviewees also mentioned the difficulty of fitting evaluations into public policy cycles, as many governors plan their strategies according to their mandates cycles, which can vary from four to five years. This means that evaluations may not be conducted frequently enough or may be conducted too late to inform policy decisions. In summary, institutionalization models in countries challenge evaluations by prioritizing economic outcomes over social or environmental impacts, using methods that don’t capture the voices and
perspectives of the groups being evaluated, and making it difficult to fit evaluations into public policy cycles. Short public policy cycles can pose a challenge to transformative evaluations because they may not provide sufficient time to assess the long-term effects of policies.

Competencies must be developed and enhanced in response to the needs of a world in transformation. Interviewees pointed out several challenges, starting with the need for systemic, holistic, and contextual responsive capacity building in the countries. At the country level, there is a need to establish competencies for evaluations and proper individual capacities. Finding education programs that fit personal limits and needs is also a challenge. Individual capacity building should focus specifically on two key elements: the generation of value judgments and the questions that guide evaluations and avoiding becoming too entrenched in methodological discussions. The interviewees explained that challenges faced by African countries differ from those faced by more developed countries, and fundamental issues such as the availability of data systems and proper budgeting for public policies evaluation have still to be addressed. Before that, institutional challenges like attracting quality staff, having efficient processes, using technology, and valuing processes that work are also basic issues. After all, interviewees pointed out that there is a lack of resources, time, and political will from implementing and donor organizations to allocate funds for evaluation capacity building. So, establishing systemic, holistic, and contextually responsive capacity building in countries requires formulating competencies for evaluations at individual and institutional levels according to these concepts instead of being mere repetitions of competencies already formulated in more developed countries. At the individual level, it is crucial that the countries offer contextually responsive education programs that fit personal limits and needs. It is crucial that these education programs focus on key elements of capacity such as the generation of value judgments and questions that
guide evaluations while avoiding becoming too entrenched in methodological discussions. At the institutional level, there is a need to build the basis for country evaluation systems including building data systems, establishing evaluation policies, growing technological infrastructure, and allocating proper budgeting for evaluations and for evaluation capacity building. All these capacity-building elements need to be aligned with the need to evaluate complex initiatives and systems.

One of the interviewees highlighted the dissemination of knowledge and good practices as crucial to enhance the professionalization of evaluation. Some participants highlighted common factors faced by professionals in the field that can hinder the effective sharing of knowledge. The dissemination of knowledge and good practices might be hampered by the lack of effective communication strategies. An example is the use of technical language or jargon making it difficult for non-experts to understand, which can limit the audience that can benefit from the dissemination. Another challenge is related to the culture of organizations and the broader social context in which they operate. Some organizations may be resistant to exposing evaluations of their initiative, even if focused only on the approaches and methods used. So, because evaluation is an applied science, the systematization of experiences and of research on evaluation results is an important part of the body of knowledge to be accessed by evaluation professionals. It is crucial that the evaluation community keeps promoting reflections on current practices and the potential changes that should be implemented in their work, and even though the need for transformation in the evaluation practice has been widely discussed in the community, this conversation must continue as the transformation process is long-term. Moreover, the dissemination of experiences involving the use of systemic thinking, complexity, new topics, and new approaches and methods can contribute to transformative change.
Strategies to Overcome Challenges, Barriers, or Obstacles that May Impact the Effectiveness of the Transformation of Evaluation. Based on the previous discussions, several strategies must be adopted to overcome challenges, barriers, or obstacles that may impact the effectiveness of the changes in evaluation practice to better contribute to transformative changes. Regarding the need to address equity and climate change, interviewees pointed out that the evaluation community has advocated for the inclusion of equity, social justice, and sustainability principles in evaluations, arguing that these aspects are necessary for addressing systemic change. In addition, the evaluation community has promoted the adoption of people-centered approaches to evaluation, which consider the perspectives of individuals from diverse contexts. According to them, these approaches have increased emphasis on incorporating indigenous knowledge systems in some regions, such as Africa and New Zealand, which focus on the connection between the living and nonliving. The evaluation community has also discussed the need for re-examining paradigms and assumptions that underlie the field, including those related to what is considered good, bad, valid, and rigorous. By advocating for an equity lens in evaluations and people-centered approaches, evaluators can help to ensure that policies and programs are designed and implemented in a way that considers the needs and perspectives of all members of society, particularly those who have historically been excluded or marginalized. By advocating for an environmental sustainability lens in evaluations, evaluators can help to ensure that initiatives are designed and implemented in a way that considers the potential environmental impacts, and that promote sustainable and resilient communities. This advocacy role can help to bring awareness to all the stakeholders that evaluations have the potential to contribute to transformative change in these areas, but only if they are designed and conducted in a way that considers the complex and nuanced nature of these issues.
Regarding strategies to incorporate systems thinking and new approaches in evaluations, interviewees pointed out that ongoing discussions have shown that using traditional research methods in evaluations may not always be possible or ethical, especially when resources are limited, and the focus is on outcome measurement and improvement. They report, for example, that randomized controlled settings, even in medical models, are increasingly being questioned for their limited representation of the population and their limitations in providing practice-based evidence. In addition, the specific context and community being studied can vary in different regions, and traditional research methods may not always contribute to the validity of the results. In addition, randomized control trials may not always be appropriate, given the available resources. They argue that defining research methods should be based on the research questions and not the other way around. One of the interviewees explains that this dominant paradigm influences the selection of data collection methods in evaluation, which tends to favor quantitative methods over qualitative methods due to their perceived ability to provide more reliable evidence. He adds this belief is rooted in early evaluation theories and the prominence of disciplines such as economics. He cited the book "The Structure of Scientific Revolution" by Thomas Kuhns to support the idea that every discipline is based on a set of paradigms and that debates in evaluation are often fundamentally based on differences in what is considered evidence. The interviewee also reported that his work is currently focused on process evaluations, which seek to bring together qualitative data on the functioning of a project with quantitative impact evidence. Due to economists’ skepticism of qualitative data, there are challenges in combining quantitative and qualitative data. The interviewee also discussed the ideological perspective of many economists who believe that only hard data from sources such as randomized controlled trials and big data constitutes objective knowledge. The interviewee finds
it interesting to work with economists who bring their values into their work and that the discussions about different concepts of human rights, justice, and knowledge highlight the diversity of perspectives in evaluation. So, evaluators have been critical of the conventional methods used in evaluations, specifically the preference for randomized controlled trial designs given its limitations in providing valid evidence, in opposition to several groups’ beliefs. “The principle that the selection of methods should be based on the evaluation questions and that evaluations should focus on the process identifying reasons for not achieving desired outcomes has the potential to influence the way evaluations are designed and utilized.” This flexibility and openness can potentially improve evaluation relevance for transformative change.

The evaluation community has taken steps toward creating an environment in which evaluation can thrive. One of the interviewees reported that several VOPEs across the world have been working on formulating standards, guidelines, and competencies for evaluation professionals, including evaluators, commissioners, managers, and coordinators. In addition, VOPEs have promoted conferences and dialogues with stakeholders, including academia, to exchange knowledge and expertise and stimulate research. Most of the initiatives approach equity, social justice, and sustainability. One of the interviewees highlighted the importance of including a broader range of stakeholders in these initiatives, such as policymakers, decision-makers, and parliamentarians to demonstrate the potential impact of evaluation on policy and decision-making that reduce inequalities. Interviewees from countries in the Global South provided examples of the impact of these initiatives. In South Africa the government adopted the African evaluation guidelines, and in India, the evaluation community signed a Memorandum of Understanding with the Monitoring and Evaluation Office of the government to provide technical support for evaluations. One interviewee from an international VOPE emphasized the
efforts also made by the evaluation community to increase the relevance of evaluations by influencing governments to adopt evaluation policies. She stressed the importance of dialogue between parliamentarians and the government to identify the need for evaluations in response to the global crisis. From her perspective, international and bilateral organizations can facilitate this dialogue. Another interviewee noted that while evaluators are convinced of the significance of evaluation in the public policy process, the public policy world may not necessarily share the same conviction. This lack of recognition is particularly challenging in developing countries facing financial constraints and structural adjustments. One Caribbean-based interviewee reported that finance is the primary constraint for conducting evaluations in the region and suggested that setting aside funds in the annual budget for evaluations and developing national statistical systems would demonstrate a commitment to evaluation. So, initiatives to advocate for the professional practice of evaluation can help ensure that evaluation conducted with rigor and accountability can not only be helpful for any decision-making process but, at the same time, can inoculate principles and values related to equity, sustainability, and social justice into the evaluation practice. Furthermore, this movement can stimulate the development of more equitable and just policies and programs, leading to more meaningful and lasting transformative change.

The topic of professionalization within the field of evaluation was also discussed by some interviewees. Due to the lack of recognition of evaluation as a separate discipline, there are no professional associations whose role is to regulate and oversee evaluation practices. As a result, the responsibility of establishing standards, competencies, ethical guidelines, and the like have fallen on VOPEs. Acknowledging the fact that the field of evaluation is relatively new and that it faces several challenges and issues at the country level is essential to ensure its relevance. One of
the interviewees, a Caribbean professor, suggests that leveraging evaluation culture requires further progress at the country level. First, she reports that there are still ongoing debates and misunderstandings about what constitutes evaluation and what kind of professionals or organizations can conduct quality evaluations. Another interviewee, a practitioner from Oceania, highlights situations of organizations seeking evaluation services from management consulting or auditing companies, believing that recognized brands are automatically capable of producing good results even if evaluation is not their core specialty.

I see situations where organizations don’t even get evaluators to come and do the evaluation. They’ll get an accounting or a management consulting firm to come and do it because then you’ll say, well, it’s XYZ (company), it must be good. And sometimes it is very good. And sometimes, it’s just giving them the report they want. And really high cost, not just the cost they pay, but all the cost of managing the contract, and all everyone’s time to go and tell them the stuff they already know. (PR3-OCE).

Changes will have to be also adopted on the demand side. As the interviewee noted, many Requests for Proposals fail to adequately balance the scope of work with the available resources. “Still, evaluators respond to these demands, do some key informant interviews, and write it up. And that’s their evaluation. And it’s just shameful that that’s what we think it is” (PR3-OCE).

So, commissioners must be very strategic when preparing their terms of reference and focus on key evaluation questions because they can influence the amount of time and financial resources spent on evaluation processes.

Another interviewee stressed the importance of understanding the difference between evaluation and research, with some individuals with research backgrounds having misconceptions about the nature of evaluation. A practitioner from the U.S. expressed frustration
with individuals claiming to be evaluators but lacking the expertise and knowledge compared to those with degrees from esteemed institutions. The interviewee also stressed the importance of differentiating evaluation and assessment, as confusion is common in many education and evaluation courses. So, an important step to make the evaluation field relevant is to engage in ongoing discussions to distinguish it from other areas such as research, auditing, and project management because each area serves a distinct purpose and has unique methods and tools that are suited to different contexts and goals. Research typically involves the systematic investigation of a topic or problem with the aim of generating new knowledge or understanding. Auditing, on the other hand, involves the independent examination of financial or other records to ensure compliance compared with established standards or regulations. Finally, project management encompasses the planning, implementation, and monitoring of specific activities or projects, with the goal of achieving specific outcomes or objectives. By distinguishing evaluation from those other areas, the evaluation community can clarify that evaluations are designed and conducted using appropriate methods and tools, and that their findings and recommendations are relevant and actionable for the intended audience. The main focus of evaluations is to assess the merit and worth of the subject being evaluated, with the aim of supporting decisions for improvement or change. Evaluations are conducted within a specific timeframe and with a specific audience in mind, who can be involved in the process at different levels. Evaluations use various criteria to assess performance and impact, which may include financial or other records. While evaluations may involve project management elements, they are more comprehensive and involve assessing both the process and outcomes of implementation. Research, on the other hand, has a more general focus and is not necessarily tied to a specific audience or timeline. This is particularly important when considering evaluation’s role in contributing to transformative
change, as evaluations must be designed and conducted in a way that maximizes their potential to bring about meaningful and sustainable change.

To overcome political resistance, evaluators have recognized the significance of establishing trust relationships with clients. According to some interviewees, trust relationships in evaluations are built by demonstrating their position as allies in improving initiatives and providing credible evidence to support informed decisions. One respondent believes this approach does not compromise her impartiality as an evaluator. Instead, she believes that individuals involved in the program have the best knowledge and that evaluation should provide a safe space for collective reflection and improvement. Another strategy mentioned as useful by the interviewees is having open conversations with commissioners about expectations and resources, assisting them in making informed decisions about how to allocate their limited resources. However, even when building trustworthy relationships, evaluation can sometimes lead to negative conclusions. This can frustrate stakeholders, leading to resistance to evaluation within organizations, particularly in the public policy arena where evaluators may be seen as policing agents. These negative feelings challenge evaluators, who are sometimes obligated to present negative conclusions to clients who have hired them. Therefore, it is essential to have these conversations and to identify and lift practices that facilitate this dynamic, considering the organization’s specific circumstances. So, the interviews show that evaluators have already realized that political resistance to evaluations is largely caused by feelings of anxiety. Those responsible for the programs feel threatened by the possible negative results that the evaluations may reveal. However, these feelings can be mitigated by an agreement between the parties that the purpose of the evaluations is to improve initiatives based on collaborative work. Ultimately,
when both parties share the same values regarding equity, social justice, and environmental sustainability, evaluations can contribute to transformative changes.

The evaluation community has discussed the need to reassess or enhance evaluation competencies and capacities to tackle new challenges. From some interviewees’ perspectives, evaluators must collaborate and adopt a more interdisciplinary approach. They believe that historically, evaluators have sought to become experts in methodologies, believing they did not need to possess extensive knowledge of subject areas and that specialists on their teams could fill any gaps. However, one of the interviewees points out it is imperative to comprehend particular sciences, such as environmental science and the interactions between human and natural systems. The interviewee stated, "We need to work with multidisciplinary experts, so they can start to add value to each other’s work. For example, I will never become an ecologist, but I need to understand how ecosystems work to provide informed advice from my specialization." (AC4-AFR). Another area cited as requiring improvement is data analysis. Although data-driven approaches to evaluation offer quicker access to evidence, they do not necessarily facilitate sound decision-making, as it requires the ability to make qualitative assessments of value and worth, which are challenging skills to develop. Another interviewee emphasized the importance of evaluators having a solid understanding of the core aspects of evaluation, including the methods and processes involved in attaching value judgments to something based on evidence. So, even though not all aspects of evaluation competencies were addressed in the interviews, it became clear that to expand the evaluation practices for new topics and to incorporate new approaches and methods, the way competencies are formulated will also have to change. The understanding of what evaluation professionals must know, and the abilities and attitudes they
must acquire to tackle new challenges of a world in transformation, will represent an important step to contribute to transformative changes.

**Implications for Evaluation Education.** Formal evaluation education has an important role in the professionalization of evaluation. One interviewee stated that, despite the lack of regulation in the field, it is crucial to be exposed to a core curriculum to be competent in evaluation. Another interviewee added that, from his point of view, to be competent in the field, previous experience and maturity are also needed beyond simply having a degree. A former administrator of an evaluation education program in Oceania also emphasized the importance of ongoing learning and self-awareness to understand the strengths, limits, and biases, and mapping social capital within the evaluator’s network, as crucial factors for professionalization in the field. So, according to the interviewees, formal evaluation education is a necessary condition, but not enough to form competent professionals in evaluation. To be able to perform evaluations for transformative change, professionals must be aware of the contemporary knowledge and abilities needed, be able to self-assess their competencies, and analyze the adequacy of formal evaluation education programs. Previous experience and maturity are also crucial elements.

During the interviews, diverse viewpoints on the potential design of evaluation education programs aimed at equipping professionals with the requisite skills to navigate evaluations amidst a rapidly evolving world were provided. One interviewee from India believes evaluators should have an interdisciplinary background and learn about different perspectives and what kind of systems change is required to address the significant issues currently threatening the planet. Another interviewee, who works for a learning center in Africa, believes that these programs should have specific foundational knowledge and not have a standardized curriculum, as evaluators have different needs depending on their areas and professional paths. A former head
of an evaluation education program in Oceania thinks that basic statistics, research methods, and evaluation theories should be included in a core curriculum. Another interviewee, a leader of an international VOPE, thinks evaluation education curricula should be based on competencies needed to face the challenges of our time and align with the audience’s needs. Another critical alignment mentioned is with future employers who could contribute to curricula development or updating. The interviewees also mention the need for more local ownership of curricula and less reliance on funding and donors. The diversity of perspectives shows that different criteria can be used to analyze formal evaluation education programs. So, even though important, a core curriculum is not enough to prepare professionals with the requisite skills to navigate evaluations amidst a rapidly evolving world. It is also necessary to have an interdisciplinary background, understand systems change, and be contextually responsive when formulating programs.

Because evaluation is a rapidly developing field, new approaches and theories which have emerged in recent years should be included in evaluation education programs. The interviewees emphasized the importance of incorporating new trends in evaluation education programs and the need to teach a broad range of theories and tools. Some new approaches mentioned include participatory, empowerment, equity, and gender-focused evaluation, which have gained significance considering the idea that evaluation can support social justice and environmental sustainability. Students must also learn to select appropriate approaches based on context, feasibility, and stakeholders’ interests. One practitioner highlighted the importance of learning about innovative approaches such as empowerment and participatory evaluation, which have roots in anthropology. Utilization-focused evaluation was also cited as an approach that can improve stakeholders’ engagement and the likelihood that the findings will be used by decision-makers. The decolonizing evaluation movement has also led to the expectation of new evaluation
frameworks informed by indigenous ways of knowing to be included in university curricula. The inclusion of systems thinking and transformational evaluations have also been emphasized as crucial. One interviewee from Oceania reported that a group of practitioners developed a new approach called "Footprint Evaluation," which orients how to incorporate environmental sustainability in all evaluations. Culturally responsive evaluations have also been highlighted as necessary, as context plays a crucial role in understanding the success of an initiative. The ability to translate knowledge and collect different perspectives has also been emphasized as essential in evaluation education programs. Additionally, democratic and cost-inclusive evaluations have been identified by some interviewees as important areas that have been neglected in the evaluation literature. So, even though evaluation has a vast number of approaches and theories, some of them have the potential to support social justice and environmental sustainability in a more significant manner. Consequently, they should be included as part of the curriculum in evaluation education programs.

Evaluators must learn a broad range of methods to be able to make choices according to the evaluation purpose, questions, and context. The interviewees presented various perspectives on the methods and techniques used in the field of evaluation. Most of them emphasized the need for evaluators to have a comprehensive understanding of both qualitative and quantitative methods and designs and the ability to adapt to different forms of evidence collection. Additionally, they stressed ground-truthing’s significance in evaluation processes, data interpretation, and program operations verification. One of the interviewees also emphasized the importance of incorporating big data sources into evaluation education programs and highlighted the need for collaboration between data scientists and evaluators. Another interviewee from the Caribbean emphasized the significance of learning to analyze qualitative data, often not taught in
academic programs. So, evaluation education programs must include a broad range of methods to equip evaluators with the skills necessary to design instruments, collect data, and analyze and interpret data to provide the most valid evidence and recommendations for decision-makers. To navigate evaluations amidst a rapidly evolving world, evaluators must be prepared to use quantitative and qualitative methods and new technologies such as big data.

Several complementary skills are essential for navigating evaluations amidst a rapidly changing world. These skills include, but are not limited to, incorporating systems thinking, cultural responsiveness, interpersonal abilities, process facilitation, project management, negotiation, and the ability to communicate complex ideas in evaluation. They are crucial for successful evaluation in complex environments. Both practitioners and academic experts believe these skills can be acquired through role-play, practical exposure to programs and projects evaluation, and on-the-job training. The interviewees also emphasize the importance of skills such as feedback, leadership, coordination, and team management in evaluations. The respondents suggest that evaluators should strive to develop these complementary skills to help them facilitate engagement and participation, manage different perspectives, build trustworthy relationships, and communicate meaningful insights to different groups. So, practitioners and academic experts believe that these skills can be acquired through practical exposure to programs and projects evaluation, role-play, and on-the-job training. Such experiential learning helps evaluators to apply theoretical knowledge to real-world evaluation scenarios, learn from mistakes, and develop critical thinking abilities. Furthermore, it enables evaluators to understand the complexities of different evaluation contexts and develop a holistic understanding of the factors that influence program success. By gaining these skills, evaluators are better equipped to
design and implement evaluations that are culturally sensitive, inclusive, and responsive to the needs of diverse stakeholders, making meaningful contributions to systems change.

Incorporating practical experiences in evaluation education programs is crucial to equipping professionals with the requisite skills to navigate evaluations amidst a rapidly evolving world. Most interviewees emphasized the importance of incorporating practical experiences, such as real evaluation projects or internships, in evaluation education programs. Such experiences enable students to acquire practical skills that cannot be learned through theoretical teachings alone. Moreover, students can develop their critical thinking, analytical, and communication skills through presentations, papers, and projects that receive instructor feedback. One interviewee who leads an international VOPE thinks that including a practicum component in EEPs will likely attract more individuals to these programs. Another interviewee suggested partnering with private sector organizations to offer collaboration on small-scale evaluations could benefit both parties. A practicum component in EEPs can also provide valuable connections between students and potential future employers, reducing the risk of unemployment for those who have completed their training. Additionally, mentorship from experienced professionals can be a valuable component of EEPs, providing students real-world experience of the nuances and complexities of the evaluation process. However, one interviewee pointed out that the low availability of professionals in certain regions may be a barrier to incorporating mentorship into EEPs. Opportunities to practice and to see how theories are translated in the real world are crucial to prepare competent evaluation professionals. Moreover, providing practical experiences is essential to prepare them for the changes that evaluation will face in a world in transformation. So, opportunities to apply evaluation approaches like systems thinking, culturally responsive, equity-focused and the ones aiming for a more environmentally
sustained world, can put students at the forefront and effectively prepare them to participate in evaluations aimed at transformative change.

Teaching methods in evaluation education programs must cater to professionals from diverse backgrounds and needs. One interviewee emphasized the need for a change in perspective among universities to leverage graduate students’ knowledge and experience and facilitate joint learning. Another interviewee further developed this idea, suggesting that classes should be viewed as learning communities with learning agendas incorporating instructor-led, student-led, and visitor-led learning experiences. The curriculum should be flexible and open to peer knowledge exchange, partnerships with governmental and non-governmental organizations, and self-directed learning. Students would be free to shape part of the program based on their interests, produce research, and share it with the community, which would be managed and made accessible through a curation process. A traditional teaching method of imparting information from the instructor to the students was deemed unhelpful for graduate programs in evaluation, as it ignores the relevant knowledge and experience brought by the students. So, evaluation is a field that attracts people from different backgrounds and with different experiences and interests. People looking for a postgraduate course, generally, already had some kind of experience. As evaluation is an applied science, the sharing of experiences and good practices, as mentioned earlier, could take place from the school benches. Currently, this role is performed primarily by VOPEs. However, universities that offer postgraduate courses could be an environment for learning exchanges that always result from evaluations. In this type of exchange and joint learning environment, reflections on the future and connections with actors in the field who are more connected with contemporary trends in evaluation would be valuable for training professionals.
Online learning in evaluation has pros and cons. The interviewees brought different opinions about online evaluation education. Some interviewees believe that online programs can be beneficial for students struggling to balance their personal lives, jobs, and studies, as it reduces barriers to accessibility. However, the quality and assessment requirements of online programs must be met. Conversely, some interviewees believe that online programs cannot provide a good learning experience and may not be suitable for all students as they miss the benefits of face-to-face learning. In addition, some students may not adapt well to the online environment, and foundational evaluation skills are best acquired through on-the-job learning.

An alternative option, a hybrid/blended modality, was also mentioned by another interviewee. One interviewee, the coordinator of a blended program in Europe, believes that a mixture of both online and face-to-face education is the best choice. Online education is more flexible for students to combine with their professional and private life, while face-to-face education allows for personal interaction and building relationships. The interviewee’s program requires students to attend in-person sessions five times over two years, which they believe is a good balance of both modalities. So, the most recommended modality for evaluation education is a mix between face-to-face and online experiences. On one hand, face-to-face experiences allow for interactive discussions, leading to a deeper understanding of evaluation concepts, methods, and techniques. It also enables students to develop interpersonal skills, essential for success in the evaluation field. Additionally, students can observe and learn from experienced evaluators, gaining a more nuanced understanding of the practical challenges and factors influencing program success. In-person interactions with instructors and classmates, along with observing evaluators in action, can reinforce learning, enhance critical thinking skills, and provide valuable insights for future evaluation work. On the other hand, online modalities can enable access to education programs
and accommodate personal contexts. Students can access course materials and lectures at their own pace and convenience, which can be particularly beneficial for working professionals who have other commitments. Additionally, online learning enables students to connect with instructors and classmates from diverse geographical locations and cultural backgrounds, leading to a more inclusive and global learning experience. Online learning also provides opportunities for students to develop digital literacy skills, which are becoming increasingly important in the evaluation field. Overall, online learning offers many benefits for evaluation students, making it an important mode of education in today’s rapidly changing digital landscape.

Evaluation education programs must prepare professionals from diverse sectors to assume different roles in evaluation. Some interviewees discussed the target audience for graduate programs in evaluation. One interviewee believes that these programs should acknowledge the differing training needs of various roles in evaluation and that external evaluators, scholars, and commissioners, for example, should be prepared differently. Another interviewee suggested that more people from the private sector and other sectors should be included in these programs. During the interview, a former head of an evaluation education program expressed her belief that admitting inexperienced students to evaluation programs would not significantly contribute to the learning environment. Instead, she argued that individuals with professional training and work experience would bring more value to the programs. Effective evaluations require input from experts in different domains, such as program design, data collection, data analysis, and the communication of findings. Additionally, evaluations need to consider the perspectives of various stakeholders, such as program beneficiaries, funders, and policymakers, to be inclusive and responsive to their needs. By preparing professionals from diverse sectors to assume different roles in evaluation, evaluation
education programs can foster a more collaborative and integrated approach to evaluation. This can lead to evaluations that are more comprehensive, rigorous, and relevant to the needs of stakeholders. Furthermore, preparing professionals from diverse sectors to assume different roles in evaluation can enhance their capacity to work together effectively and to appreciate the value of different perspectives and expertise. This, in turn, can lead to improved communication and decision-making, which are critical for the success of evaluation efforts. This will ultimately lead to more effective evaluations and improved outcomes for programs and interventions.

One of the interviewees described the assessment system of a graduate program she coordinated in Australia, where students demonstrated their learning by applying evaluation techniques and approaches, adding self-reflection, and documenting their work in a portfolio or learning log. The interviewee noted that the shift to this assessment system was challenging but had the potential to turn into a knowledge-sharing platform that could add practical knowledge to the academic literature. The interviewee envisioned a community of learners that share their distillations and engage in continuous improvement. The interviewee suggested this kind of assessment system would be revolutionary in evaluation education. So, the academic environment should stimulate knowledge sharing in evaluation, as it is an important element of learning and professionalization. A knowledge-sharing platform about the topics, approaches, and methods relevant to systemic change could add practical knowledge to the academic literature in this aspect.

Evaluation education programs should include a professional component. One interviewee reported an ongoing debate about evaluation professionalization. She believes evaluation education programs should include topics related to evaluation professionalization and the benefits of being a member of organizations focused on developing standards and ethics.
in evaluation. Other interviewees also emphasized the need for educational programs to have a professional component that supports students’ preparation for the labor market by providing orientation on how to apply for evaluation positions or responding to Requests for Proposals. So, including a professional component in evaluation education programs can allow young evaluators access to standards, ethical guidelines, and competencies related to the professional practice of evaluation. Also, by participating in evaluation conferences promoted by VOPEs or other organizations, they network and attend discussions on new trends in the field by experienced professionals and access new literature. By being in touch with new trends, professionals can expand their perspectives and understand why and how evaluation can play a role in contributing to transformative change.

**Conclusion**

From the interviewees’ perspectives on the meaning of systemic and transformative change, systemic change is an overarching and interrelated change that alters existing paradigms, affects power relations, and brings previously unseen actors to light. Transformative change refers to bringing excluded groups to the table, reshaping the system, and promoting transformative learning. The interviewees agree that evaluation has the potential to contribute to transformative change by providing information and evidence to understand problems, design solutions, measure progress, and adjust strategies. However, for evaluations to effectively contribute to transformative change at the country level, they must consider larger societal issues, such as equity and environmental sustainability, and be conducted within national evaluation systems, strategically positioned in the countries’ governmental organizational structure. They also emphasized the importance of conducting evaluations with a more systemic view to have a more significant potential to contribute to transformation.
Current evaluation practices are not effectively driving transformative and systemic change, and a repositioning of evaluation practice is necessary to fulfill evaluation’s potential for contributing to social change and justice. In general, the changes in evaluation practice include a greater emphasis on better theories of change, responsiveness to context, attention to equity and environmental sustainability issues, and the inclusion of the perspectives of those who have been previously unheard. The current dominant evaluation models, centered on program evaluations, are linear and one-dimensional and do not reflect the complexity of interventions. Incorporating systemic thinking into evaluations can change the purposes that have guided evaluations in most institutional settings and shift the focus towards collaboration and contribution, promoting lasting change.

The change in evaluation practices also encompasses building up evaluation ecosystems in regions and countries still in the initial stages of evaluating their initiatives. Building evaluation ecosystems requires continuous training, technical assistance, research, policy development, and articulation among stakeholders such as higher education institutions, private consulting firms, statistical agencies, and government agencies. The objective is to improve the overall efficacy and ownership of evaluation processes.

Evaluation is a rapidly evolving field that has traditionally been used in social settings but has since expanded to other themes such as strategies and organizations. According to the respondents, the future of evaluation is likely to focus on several key topics, including sustainable development and climate change, indigenous knowledge systems, conflicts and their economic and social impacts, social justice and human rights, workforce changes, and the effects of the COVID-19 pandemic. The adoption of these topics in evaluation practice will be influenced by institutional culture, with some organizations being more open to adopting specific
lenses such as equity and sustainability. However, there is a discrepancy between the scholarly work on evaluation, which highlights the importance of these topics, and the actual practice of evaluation in large systems within institutions such as the European Commission, multilateral development banks, and national governments. Finally, technology has become an important topic in evaluation due to its ability to save time and money, enable collaboration and exchange of knowledge, facilitate lifelong learning, and overcome language barriers. These benefits are crucial for evaluators who need to keep up with the constantly changing landscape of their field.

The interviewees mentioned seven types of approaches they believe will gain importance in the field of evaluation in a world in transformation. First, the systems-oriented approach emphasizes the importance of incorporating complexity and systems thinking in evaluations to understand and bring about transformational change. Second, the empowering, participatory, or deliberative approaches aim to provide a voice to different groups and involve them in the evaluation process, leading to more meaningful and relevant findings. Third, the developmental evaluation approach promotes iterative, adaptive processes and continuous feedback to drive transformative and systemic change. Fourth, the contextually responsive evaluation approach reorients evaluations to be responsive to the local contexts. Fifth, equity-focused evaluation approaches highlight the importance of focusing on marginalized and excluded groups. Sixth, the utilization-focused evaluation approach is particularly important for transformative change because it emphasizes the importance of stakeholder engagement and participation and focuses on the use of evaluation findings to inform decision-making processes. Finally, approaches like “Footprint Evaluation” can orient how to incorporate environmental issues in evaluation processes.
The COVID-19 pandemic led to a rethinking of traditional evaluation methods, with a growing consensus that the exclusive use of quantitative methods has limitations. For example, such methods may not capture critical subjective experiences or perspectives and may not be suitable for complex or multi-faceted programs. Big data, for example, can speed up evidence collection but may have biases and limitations. To address these limitations, some experts advocate using qualitative or mixed methods. Qualitative methods can ensure representation and a voice for different groups in the evaluation processes. The principle of ground truthing, which involves verifying the validity of data interpretation, highlights the importance of direct observation and engagement in fieldwork to gain a comprehensive understanding of community experiences and needs.

The field of evaluation will need to undergo significant changes to adapt to the challenges of the current era. The demand and supply sides of the field will need to revise their competencies and capacities in response to these new challenges. On the demand side, terms of reference will have to be prepared with more strategic considerations to ensure the efficient use of time and financial resources. On the supply side, evaluators will have to embrace multidisciplinary collaboration and expand their understanding of thematic areas such as climate change and environmental sustainability. They will also need to develop their data analysis skills and the ability to make evaluative judgments based on qualitative data. Evaluators must also work to establish policy dialogues with governments and parliamentarians to demonstrate the value and relevance of their work.

Although evaluation can be a powerful tool for promoting transformative change, its potential is not widely recognized in many parts of the world, particularly in the Global South. The perception of evaluation as a means of driving systemic change varies among individuals
and organizations. Some interviewees reported advances in their local contexts through increased attention to implementation learning, the connection of programs and stakeholders, and meaningful practice of evaluation. However, there are still challenges in incorporating sustainability and equity considerations into evaluative practices, focusing on marginalized populations, and incorporating complexity into evaluations. Some interviewees emphasized that equity should be a cross-cutting criterion in evaluation. Agenda 2030 and the 17 SDGs are seen as opportunities to stimulate systems thinking in evaluations.

Various stakeholders discussed challenges facing the evaluation industry through interviews. The first challenge highlighted is the need to make evaluative evidence helpful and relevant in public policy, which requires well-designed evaluation functions and policy dialogues with governments and parliamentarians. Another challenge is establishing evaluation as a transdisciplinary field, requiring evaluators to have cross-disciplinary skills, and be closely connected to other professional communities. Building up evaluation ecosystems and addressing equity and climate change are also noted as challenges. The influence of institutionalization models, donors, and political aspects on the evaluation process is discussed, as well as the need for evaluators to balance their values with pragmatism. Concerns about the focus and approaches of evaluations, including the limitations of OECD evaluation criteria, were also highlighted.

The field of evaluation is still in its early stages and faces numerous challenges and barriers. There is a need for further progress in developing the evaluation culture and distinguishing it from other areas, such as research, auditing, and project management. The evaluation community strives to overcome resistance by establishing trust with clients and having open conversations about expectations and resources. Efforts are also being made to professionalize the field by creating standards, ethical guidelines, and competencies for
evaluation professionals, including evaluators, commissioners, managers, and coordinators. The evaluation community is also promoting a people-centered approach to evaluation, which considers individuals’ perspectives from diverse contexts and emphasizes the principles of equity, social justice, and sustainability. However, the political power present in evaluation processes can pose a barrier to achieving these goals, as decisions regarding evaluations are often made by funders and donors rather than by local governments, leading to a lack of local buy-in and a disconnect between the data collected and the needs of the local community.

Interviewees presented their perspectives on the role of formal training and education in becoming a professional evaluator. They discussed different aspects of evaluation education programs, including the importance of foundational knowledge, ongoing learning, interdisciplinary background, practical experiences, and complementary skills. The perspectives on the methods and techniques used in the field of evaluation emphasized the need for a comprehensive understanding of both qualitative and quantitative methods, ground truthing, and incorporating big data sources. The interviewees also highlighted the significance of practical experiences in evaluation education programs and the value of mentorship from experienced professionals. The upcoming session will feature the findings from the focus group conducted with young and emerging evaluators.

**Study 2: Focus Group with Young and Emerging Evaluators (YEEs)**

Study 2 aimed to collect the perspectives on the role of evaluation to support systems changes and interests and needs regarding evaluation education from a group of professionals representing the field’s future workforce: the young and emerging evaluators (YEEs). According
to the Evalpartners initiative\textsuperscript{7}, YEEs are evaluators under 35 years or professionals who have less than five years of experience and have expressed an interest in becoming professional evaluators within the past five years or less (Evalpartners, 2016). In 2015, EvalYouth Global, a global network of YEEs, was launched and supported by Evalpartners initiative. Since then, this global network has grown and developed, leading to regional and national chapters. The number of current EvalYouth Global followers is estimated at 8,700 YEEs. Following this, the analysis findings will be presented, utilizing a coding system consisting of ‘YEE’ followed by the letter P and a sequenced numeral to designate YEEs focus group participants.

\textbf{Findings}

\textbf{Participants' Understanding of the Term Transformational Change.} For the YEEs who participated in the Focus Group, transformational change means a paradigm shift in the way people see the world moving from traditional linear models to a more systemic way of thinking. They pointed out that the COVID-19 pandemic is an excellent example of how everything is connected, and how important it is that evaluation recognizes that.

For me, the (COVID-19) pandemic is a very good example because, at least from my perspective, when everything started, it was in China, far away from my country, and we never thought it could have such a huge impact. And now, for example, with everything happening in Ukraine and Russia, it’s like, oh, it’s far away; we don’t even have the same language or anything. But still, it has an impact on our lives. And I think that transformational evaluation and transformational change, for me, is to recognize that everything is connected, and we need to switch the way we think about the problems. (YEE-P1).

\textsuperscript{7} https://evalpartners.org/
They added that policies, and consequently evaluation, should focus, for example, on how local events can impact the global level and vice versa. Hence, effecting a transformational change necessitates a comprehensive perspective concerning the constituents of problems and their interdependencies. Consequently, to foster transformative change, evaluators need to alter their approach to evaluating objects by acknowledging the interdependent nature of their constituents.

Transformational change requires system thinking. The participants point out that to increment the systemic view, evaluators need to seek knowledge in a broader range of areas such as climate change and the 2030 Agenda for Sustainable Development. They note that current discussions on topics such as climate change tend to focus too much on environmental specialists rather than evaluators. They also emphasize that to think systemically about issues that are affecting society, there is a need to include those who have been left behind, particularly in Africa, where evaluation has traditionally been viewed as a "white man’s affair." To cultivate a systemic perspective on intricate issues such as climate change, sustainable development, and poverty, evaluators must expand their knowledge beyond conventional areas of expertise in evaluation. This expansion of knowledge will enable evaluators to engage in discussions concerning these issues more effectively and enhance the relationship between them and the evaluation field.

Themes or Topics Likely to Gain Importance for Evaluation Practice in the Coming Years. Several topics are likely to gain importance in the evaluation practice in the coming years. Elevating the activity of evaluation to the status of a profession will be essential for the field to achieve the necessary relevance to contribute to the major challenges of the future. One key topic mentioned by the participants is the professionalization of evaluation. From the
participants’ perspective, there is a need to establish a firm grounding for evaluation and to recognize it as a career or a profession, just like law or medicine. Again, they mention that competency frameworks are needed as references for a more professional evaluation and would build credibility and help in the professionalization of evaluation. By establishing a clear career path for evaluators, evaluation can gain the necessary recognition and credibility to attract and retain top talent.

Young people have an important role in shaping the future and achieving sustainable development. YEEs think youth is another topic that is likely to gain more importance in the evaluation practice in the coming years. They point out that most of the global agendas for sustainable development, such as the Agenda 2030, have put youth at the center of all development issues. As the world’s largest youth population in history, young people are key stakeholders in achieving the UN Sustainable Development Goals (SDGs) and creating a more equitable and sustainable world. Furthermore, youth are disproportionately affected by many of the global challenges, such as poverty, inequality, unemployment, and climate change, and thus their inclusion in decision-making processes is critical for finding effective solutions. Recognizing the potential of youth, international agendas are increasingly prioritizing youth engagement, empowerment, and participation in policymaking and implementation processes to create more inclusive and participatory societies. The participants also stressed that including youth in evaluation processes that target them and the issues related to them, like employment, will become more critical. Also, there is a need to help YEEs become professionals and participate effectively in evaluation endeavors in general. The YEEs’ engagement is important for several reasons. Firstly, it ensures that youth perspectives are heard, providing valuable input into decision-making. Secondly, involving YEEs can build evaluation capacity, diversifying the
field of evaluation and bringing new perspectives. Finally, it empowers youth, allowing them to take an active role in shaping their own futures and building leadership skills.

Decolonization of evaluation is becoming an important topic. The discussion among YEEs regarding evaluation decolonization suggests that a systemic approach to evaluation is needed, with a focus on creating evaluative thinking that is context-responsive and people-centered. The discussion highlights the importance of considering specific topics such as gender, climate change, and inequalities, not as separate courses, but as a way of thinking. Decolonizing evaluation is seen as crucial in addressing these issues, with a need for a more rigorous approach to methods that is context-sensitive and responsive to the needs of marginalized groups, including youth, women, and indigenous peoples. Overall, the discussion emphasizes the importance of centering people in evaluation processes and promoting a more equitable and inclusive approach. By decolonizing evaluation, dominant power structures and ways of thinking that have historically marginalized and excluded certain groups, particularly indigenous peoples, people of color, and other marginalized communities, will be challenged. By centering the voices and experiences of these groups in evaluation processes, decolonizing evaluation can help to create more equitable and just outcomes. It can also help to shift the focus from a narrow understanding of changes based on linear models to a more holistic focus. Also, it can incorporate contextually sensitive approaches that consider multiple forms of knowledge and ways of knowing. Ultimately, decolonizing evaluation can contribute to transformative change by promoting a more inclusive, equitable, and just society.

In the future, technology is likely to play a crucial role in evaluation practices. The focus group participants, YEEs, stressed the critical role of technology in evaluation practices going forward. They point out that integrating technological resources into evaluation processes and
keeping up with advancements is vital for their profession. Technological utilization by individuals and organizations has made significant progress over the past three decades since the public gained access to the Internet. It has brought about changes in how people interact, expanded possibilities for engaging with diverse groups in different locations, and simplified and sped up data collection, organization, and analysis. Additionally, communication tools have facilitated outreach to diverse audiences and enabled data visualization. In addition, access to knowledge and information was also facilitated, contributing to evaluation professionalization. Failing to embrace new technological resources could render organizations and evaluators obsolete and less competitive. Staying informed and making informed decisions about how to adapt to changing circumstances is possible with technological advancements. Therefore, incorporating technological resources into evaluation practices and keeping up with technological transformations is crucial for remaining effective and competitive in an ever-changing technological landscape.

Overall, in the coming years, participants anticipate that the evaluation practice will prioritize several topics. One key topic is the professionalization of evaluation, which involves establishing a clear career path for evaluators to attract and retain top talent. Another important topic is youth engagement in evaluation processes, as young people are key stakeholders in achieving global agendas, such as the UN Agenda for Sustainable Development, and building more equitable and sustainable societies. Decolonizing evaluation is also becoming increasingly important, with a need to create evaluative thinking that is context-responsive and people-centered, particularly for marginalized groups. Finally, technology is expected to play a crucial role in evaluation practices, with a need to integrate technological resources and keep up with advancements to remain effective and competitive.
**Approaches and Methods to Facilitate Transformative Change.** Some evaluation approaches are likely to gain more importance in the coming years. One of the approaches that emerged from the discussion is Utilization-focused Evaluation, which addresses how stakeholders use evaluation findings to inform decision-making. In addition, the group emphasized approaches that ensure evaluation is people-centered, which involves incorporating the perspectives and experiences of various stakeholders, especially minorities like youth, women, and indigenous people. To achieve this, the group suggests decolonizing evaluation to make it more context-responsive and inclusive. Another aspect mentioned in the discussion was the need to use systemic evaluation approaches. The group highlighted the need to focus on systemic change and address specific topics such as gender, climate change, and inequality. Additionally, the group discussed the need for a more gender-responsive approach to evaluation that incorporates diversity and inclusion. Regarding innovative approaches, the group mentioned the potential use of foresight techniques, which analyze future scenarios, to inform evaluation practice. The group suggested that this approach could contribute to more predictive thinking about transformations that will happen in the medium and long term.

Regarding specific topics, I would say, actually systemic change and systemic approach to evaluation, but put different focuses, for example, on gender, climate change, and all inequalities. I would not suggest any specific course on climate change, but rather a way of thinking. Because, for me, evaluation is rather than focusing on the need to be rigorous in methods. But the most important part is creating evaluative thinking, so yes, I do believe we need to focus on youth, women, and indigenous. But I think decolonizing evaluation will bring us a lot of that. (YEE-P8).
The YEEs’ emphasis on these approaches indicates a rising awareness of the imperative to link evaluation goals with the systemic modifications required to tackle the planet’s pressing challenges. This also signifies an acknowledgment that evaluations must scrutinize whether initiatives are devoted to bringing about transformative changes. Such a commitment entails comprehending the contexts and genuinely listening to people’s requirements while ensuring that evaluation findings are utilized by decision-makers in a contextually appropriate and inclusive manner. Additionally, it involves a consideration of the future.

Mixed methods are likely to gain importance in evaluation practice in the coming years. The YEEs in the focus group discussed their perspectives on methods likely to gain importance in evaluation practice in the coming years. One method emphasized was qualitative research, particularly the use of storytelling techniques. Participants noted that while surveys are commonly used for data collection and graphics are currently used for data visualization, there is a growing need for more people-centered approaches that allow for a deeper understanding of the experiences and perspectives of those being evaluated. Storytelling was seen as an excellent way to achieve this, but there was also recognition that it can be challenging to analyze large amounts of qualitative data. Despite highlighting qualitative methods, the group also recognized the continued importance of quantitative methods, such as randomized approaches for comparing control and experimental groups. However, participants acknowledged that ethical issues come with this type of research, particularly regarding who is assigned to treatment groups. Overall, the group suggested that a mixed-methods approach that combines qualitative and quantitative methods would be most effective for evaluation practice. In addition, they noted the importance of using technology to analyze and make sense of large amounts of data, particularly when it comes to qualitative methods like storytelling. Finally, there was a recognition of the cultural
significance of storytelling approaches in the Global South contexts, particularly Africa, and the need to adapt evaluation methods to fit local cultural norms and values. The discussion on data collection and analysis methods was consistent with the previously identified relevant approaches by the YEEs. It is crucial for methods and approaches to align with evaluation goals and principles. By using mixed methods, evidence regarding causation and impact can be both measured and contextualized through the perspectives of stakeholders involved in the initiative being evaluated. The triangulation of these different sources of information can yield a more comprehensive and reliable set of evidence, increasing the likelihood of effective and transformative decision-making.

Incorporating participatory methods and utilizing innovative instruments for data collection are critical components in facilitating evaluation’s contribution to transformational change. When asked about what approaches would be better to contribute to transformational change, most participants mentioned participatory approaches. Those approaches can ensure that those evaluated can express their situation from their perspective, as the findings will be used to make decisions that affect them in the short and long term. Regarding methods and tools, participants suggested using new instruments for collecting data, such as storytelling and artificial intelligence (AI) will represent an important change in the way evaluations collect evidence. For them, adopting storytelling, for example, would ensure inclusivity in evaluations by reaching underrepresented groups. On the other side, the use of AI in evaluations can assist evaluators in analyzing data in a more comprehensive, accurate, and timely manner. They added that the use of AI can provide a better understanding of complex issues. Thus, the use of storytelling can increase the validity of data by incorporating real-world perspectives, and the use of AI can help evaluators to accelerate transformative change by providing innovative insights
and providing timely evidence to decision-makers. In addition, YEEs believe that evaluation must focus on learning, rather than audit or policing, to encourage better decision-making, identify what has worked and what has not, and improve initiatives in the future. So, as per the participants’ insights, the contribution of evaluation to transformative change necessitates modifications in evaluation processes, ranging from the purpose of evaluations to the approaches, methods, and techniques employed in gathering and analyzing evidence.

**Inclusion of YEEs in Evaluation Processes to Bring Novel Perspectives to Transformative Evaluations.** Incorporating YEEs in evaluation processes serves not only to enhance their individual capacities but also to bring novel perspectives to transformative evaluations. The focus group participants also expressed concerns about the lack of youth involvement in evaluations and their belief that young people’s perspectives should be included in the evaluation process. They also emphasized the importance of collaboration with other evaluators globally as crucial for developing interventions. They propose that students should have access to communities of practice to gain exposure to real-life evaluation processes and share knowledge and best practices.

… some of these groups help capacity building by sharing resources, tools, and templates. And this is how students can begin to actually get exposed to real-life evaluation processes. And so, I think that beyond the classroom work and everything, there should be opportunities in various countries for evaluation students to participate in some communities of practice, some even available online. (YEE-P1).

Therefore, incorporating young and emerging evaluators in evaluation processes can contribute to transformative evaluations because they bring fresh perspectives and innovative ideas to the evaluation process. These evaluators often have different backgrounds and
experiences that can lead to unique insights and approaches to evaluation. Additionally, involving younger evaluators in evaluation processes can help to build their capacity and develop their skills, which can contribute to a more diverse and robust evaluation workforce in the future. By incorporating young and emerging evaluators, evaluation processes can be more inclusive and reflect a broader range of perspectives, leading to more transformative and impactful evaluations.

**YEEs’ Expectations and Needs Regarding Evaluation Education.** The expectations of YEEs when it comes to evaluation education programs revolve around practical experience. Specifically, YEEs are looking for hands-on opportunities to apply evaluation techniques in real-world situations and access to tools that will allow them to participate in all stages of evaluations. In addition, YEEs are seeking short courses or programs that are accessible and inclusive, breaking down financial and physical barriers. To meet YEEs’ needs, evaluation programs should have a strong practical component, including group work, senior mentoring, and internships. YEEs believe that evaluation education should be incorporated into higher education and include research methodologies to help students understand the fine line between research and evaluation. Overall, YEEs see practical experience as a critical factor that sets young evaluators apart. YEEs seek to enhance not only their theoretical knowledge but also their abilities and attitudes. It is widely acknowledged that practical opportunities to apply theoretical concepts to real-world scenarios are the most effective means of acquiring evaluation competencies. Hence, YEEs prioritize hands-on opportunities when searching for evaluation education programs. Such practical opportunities enable them to build professional networks and gain exposure to diverse evaluation contexts. YEEs recognize that practical experience is highly valued by employers in the evaluation field, and participating in evaluation projects can increase
their chances of finding employment and advancing in their careers. Furthermore, exposure to diverse evaluation contexts can broaden YEEs’ perspectives and cultural competence, which is crucial in today’s globalized and diverse evaluation environment.

YEEs complain about not having access to evaluation-related educational opportunities that turn them into competent evaluators. The YEEs who participated in the focus group listed several challenges that they face when accessing evaluation-related educational opportunities. The first challenge was related to the lack of availability of educational programs. Some participants pointed out that there is no offer of postgraduate programs in evaluation in their countries (e.g., Mozambique, Ghana, India), which makes evaluation learning for them only possible through open sources and free online short-term courses combined with on-the-job training opportunities. Even when open sources and free online short-term courses are more accessible and affordable, they are not recognized as formal education and do not add to the curriculum. In addition, access to this kind of solution can be challenging for those who live in areas with poor internet connections or a lack of digital infrastructure. One of the participants from Latin America adds the offer of programs can be prevented in countries where infrastructure and connection costs can represent significant constraints in education initiatives.

In India, we do not have a separate evaluation education anywhere in the universities and everywhere. Moreover, people are getting training on the job. They seek opportunities on development aid programs and projects, where monitoring and evaluation is part of the funders’ policies. (YEE-P4).

The field of evaluation has failed to establish education opportunities that ensure the training of future evaluators, particularly the young and emerging ones. Consequently, when experienced evaluators retire, the number of competent professionals capable of conducting
quality and transformative evaluations is not likely to increase. This situation poses a risk of losing expertise in the field, as it cannot be easily replaced. In addition, without evaluation education access, YEEs are unlikely to be exposed to the requisite topics, approaches, methods, and tools necessary to facilitate systems change.

Evaluation Education Programs have a narrow focus preventing YEEs from fulfilling interests and needs. YEEs participants in the focus group have identified a challenge related to the narrow focus of existing evaluation education programs, which primarily concentrate on specific fields like education and psychology. This leaves individuals from other areas, such as economics, feeling excluded as they are unable to find any courses related to their educational background or interests. Depending on the admission requirements of universities, someone with an economics background, for instance, may not qualify for a master’s program in evaluation that demands a background in a particular field such as psychology. The participants think evaluation education curricula should cover multiple perspectives, which could attract a more diverse set of students. Therefore, there is a need to develop more evaluation education options that cater to a wider range of disciplinary backgrounds and perspectives. So, the limited focus of existing education programs on specific disciplines and perspectives prevents individuals from other fields from accessing relevant courses, creating a lack of disciplinary diversity, and limiting the scope of evaluation as a field. Moreover, the exclusion of individuals, particularly YEEs, from evaluation education programs can lead to a lack of inclusivity and diversity in evaluation processes, hindering the ability to conduct transformative evaluations that challenge power structures and promote equity.

Evaluation education can be too expensive for individuals who are starting their careers as evaluators. Another challenge YEEs face is the high cost of educational programs. They point
out that most of the evaluation education programs are in the Global North, which makes it costly for YEEs in the Global South to access them. They add that even summer schools or online programs, which are cheaper, are still too expensive for people from the Global South. "These courses need to be more inclusive in order to break down financial and physical barriers, especially after the pandemic when everything online has become possible and easier." (YEE-P3). So, the high cost of such programs can act as a barrier for YEEs from marginalized or low-income backgrounds, resulting in a lack of diversity and inclusivity within the field. Diversity and inclusivity are crucial for carrying out transformative evaluations that challenge power structures and promote equity. Therefore, making evaluation education programs more affordable and accessible can promote diversity and inclusivity, leading to more transformative evaluations.

Employers of YEEs should support them in pursuing formal evaluation education. According to the participants, the organizations they work for demonstrate a lack of support for their interest in pursuing educational opportunities related to evaluation. For instance, some organizations prohibit employees from taking time off for educational purposes, preventing YEEs from acquiring knowledge and developing their skills, consequently reducing their chances of career advancement. Therefore, as a vital aspect of the evaluation ecosystem, organizations that employ young laborers should not only offer practical opportunities but also encourage and support them in pursuing formal education. This will enable them to acquire theoretical knowledge, which is essential to comprehend the correlation between theories and methods and their applications in the practical field of evaluation. This understanding is crucial for conducting transformative evaluations that challenge power structures, promote equity, and better serve the needs of diverse communities. Therefore, organizations must support YEEs in
pursuing formal education as a crucial aspect of the evaluation ecosystem, promoting inclusivity and diversity in the field, and leading to more transformative evaluations.

**Evaluation Education Programs to Prepare YEEs for Systems Change Evaluations.** Practical experience is a mandatory aspect of evaluation education. This was the main topic discussed by participants as a priority for EEP. From their perspective, theoretical knowledge is insufficient, and evaluators require training in handling the practical day-to-day issues of evaluation activities. Consequently, all participants emphasized the necessity of postgraduate programs to require a practicum as an integral part of their curriculum. This practicum should offer students practical experience in evaluation processes and also the development of practical skills beyond classroom learning. Therefore, a balance between theoretical and practical aspects is imperative to meet the needs of YEEs. The practicum must entail collaboration with third-party organizations or smaller evaluation projects. The participants suggest the establishment of a system where institutions provide scholarships to enable students to work on real projects and be mentored by more experienced evaluators. The participants consider mentoring and follow-up support crucial for their training. "It is difficult to have an evaluation as a TCC and do it alone because it is usually the first time you deal with an evaluation process. It is necessary to have mentoring and follow-up as part of the training." (YEE-P7).

Furthermore, participants suggest the establishment of a practice-based learning culture within the evaluation education system. They propose that evaluation education should, as much as possible, shift away from the conventional theoretical classroom format and instead furnish students with tangible illustrations of how theory can be translated into practice, possibly in the form of practical workshops. These opportunities would allow students to develop real-life
evaluation perspectives and understandings of how evaluations are carried out beyond theoretical concepts and ideas.

The skills acquired during their education, even at the foundational levels, can ready them for future evaluation challenges such as systems change evaluations. Offering opportunities for students and YEEs would also benefit organizations in the sense that younger individuals bring fresh perspectives and can influence the use of updated technological tools. The inclusion of YEEs in evaluation processes can contribute to greater equity, diversity, and inclusion in organizational environments.

Evaluation education programs should focus not only on training consultants but professionals able to disseminate the evaluative culture and bring evaluation into the management of their institutions. The focus group participants have observed that evaluation education programs are predominantly geared towards external evaluators, neglecting other critical roles such as managers, commissioners, researchers, and teachers that can also play an integral part in the evaluation ecosystem. They acknowledge that managers and commissioners have a significant influence on the development of evaluation culture within organizations and can transform evaluation processes into learning opportunities. Therefore, it is imperative to expand the scope of the target audience in evaluation education programs deliberately to contribute to evaluation culture in organizations. Offering specific content tailored to professionals’ specific roles can contribute to strengthening the evaluation ecosystem and shift organizations’ perception of evaluation and its role in transformative change.

Evaluation education programs must be designed in a way that students can balance study and work needs. The participants highlighted flexibility as a crucial priority, considering the need to manage personal and professional lives. Among the various aspects of flexibility, one
key element involves a non-relocation requirement for post-graduation attendance. As a potential solution, remote learning was suggested. Nevertheless, despite being flexible, remote learning cannot replace the practical experience gained in the field. Therefore, the optimal approach would be to combine in-person and remote learning modalities.

Evaluation education programs should provide comprehensive instruction on research methods, statistical analysis, and various evaluation techniques and philosophies that accompany paradigm shifts. YEEs who participated in the focus group emphasized the importance of extensive training in research methodology and practices for evaluators to excel in their roles. In addition, they believe that evaluators must have a strong command of statistics, data analysis, and proficiency in tools such as R, Power BI, and Python. By pointing to curricula priorities, research methods, and data analysis, YEEs demonstrate a commitment to producing valid evidence with the use of new technologies for data collection. The focus group also recommends that evaluation education programs should teach different evaluation techniques and philosophies that accompany paradigm shifts, including empowerment and utilization-focused evaluation (UFE). Moreover, evaluators should receive training on the use of diverse tools that are applicable to various contexts and scenarios. From their perspective, to be competitive, evaluation programs should incorporate emerging trends in the field, such as the use of new technologies like artificial intelligence, unstructured data, and communication techniques. Additionally, there should be courses in languages other than English with an international standard since most of the literature is in English, and those who do not master this language end up staying on the sidelines. This would enable evaluators to gain a competitive edge and remain up to date with contemporary methods and approaches to assess systemic changes impacting the world.
Evaluation education programs should include a quality assessment parameter. Participants note that professionals often stumble into evaluation as a career, lacking a disciplinary foundation. To address this issue, guidelines, standards, and competencies frameworks have been developed to define the skills, knowledge, and attitudes necessary for quality evaluations. From the YEEs' perspective, as competencies provide a clear framework for many professional activities, formal evaluation-related education programs should be shaped by evaluation competencies to ensure that evaluators possess the necessary skills and knowledge to conduct high-quality evaluations. By adopting a competencies-based approach, universities can prepare evaluators for the demands of their roles and foster a common understanding of what it means to be a competent evaluator, ensuring consistency and quality across evaluations. When competencies encompass principles of equity, sustainability, and social justice, shaping formal evaluation-related education programs around competencies has the potential to drive transformative change.

**Engagement in communities of practice.** Lastly, the participants recommend that emerging evaluators should have the chance to partake in communities of practice, deeming it a vital element of their professional training. Communities of practice facilitate the sharing of resources and tools, exposure to real-life evaluation processes, and contribute to capacity building. Hands-on engagement is essential in preparing students for complex evaluation tasks.

**Conclusion**

YEEs believe that transformational change requires a paradigm shift from traditional linear models to systemic thinking, where everything is connected. To foster transformative change, evaluators need to change their approach to evaluating objects and acknowledge the interdependent nature of their constituents. YEEs propose that policies and evaluations should
focus on how local events can impact the global level and vice versa. They suggest that evaluators expand their knowledge beyond conventional areas of expertise to cultivate a systemic perspective on intricate issues such as climate change, sustainable development, and poverty. YEEs emphasize the importance of participatory methods and innovative instruments for data collection in facilitating evaluation’s contribution to transformative change. They also stress the need to incorporate youth perspectives in evaluation processes to bring novel perspectives and innovative ideas to the evaluation process.

The expectations of YEEs regarding evaluation education programs are centered on practical experience and hands-on opportunities to apply evaluation techniques in real-world situations and access to tools for all stages of evaluations. Practical experience is highly valued by YEEs, as it allows them to build professional networks, gain exposure to diverse evaluation contexts, and increase their chances of finding employment and advancing in their careers. Additionally, exposure to diverse evaluation contexts can broaden YEEs’ perspectives and cultural competence, which is critical in today’s globalized and diverse evaluation environment. The integration of evaluation education into higher education, the learning of research methodologies, and understanding the difference between research and evaluation are also desired by YEEs. YEEs also seek inclusive and financially accessible programs.

When asked about the priorities for evaluation education programs to meet the needs of YEEs, the participants emphasize the importance of practical experience and propose that postgraduate programs require a practicum as an integral part of their curriculum. This practicum should entail collaboration with third-party organizations or smaller evaluation projects and include mentoring and follow-up support. The participants suggest the establishment of a practice-based learning culture within the evaluation education system, offering opportunities for
students to develop real-life evaluation perspectives and understanding. They also recommend expanding the scope of the target audience in evaluation education programs to include professionals other than external evaluators, such as managers, commissioners, researchers, and teachers. Additionally, evaluation education programs should provide comprehensive instruction on research methods, statistical analysis, and various evaluation techniques and philosophies, including emerging trends in the field. Finally, the programs should be designed in a way that allows students to balance study and work needs and be flexible, considering the need to manage personal and professional lives.

The YEEs emphasized several key evaluation approaches, including Utilization-focused Evaluation, people-centered approaches that incorporate diverse stakeholders’ perspectives, and culturally responsive and systemic evaluation approaches. The group also suggested the potential use of foresight techniques to predict future transformations. Moreover, the focus group discussed the importance of creating evaluative thinking as a condition for transformative change. The YEEs recognized the significance of systemic change and a systemic approach to evaluation, with a focus on specific topics such as gender, climate change, and inequalities. The group’s emphasis on these approaches reflects an increasing recognition of the need to link evaluation goals with the systemic changes necessary to address pressing issues, while integrating diverse perspectives and a vision of the future.

The participants discussed the growing importance of mixed methods in evaluation practice, with a particular emphasis on the significance of qualitative research and storytelling techniques. They recognized the cultural significance of storytelling approaches in the Global South and emphasized the need to adapt evaluation methods to fit local cultural norms and values. While also acknowledging the continued importance of quantitative methods, they
highlighted the need to address ethical issues related to randomized controlled trial designs. The group suggested that a mixed-methods approach, combining both qualitative and quantitative methods, would be most effective for evaluation practice. Overall, they stressed the importance of aligning methods and approaches with evaluation goals and principles to increase the likelihood of effective decision-making.

**Study 3: Evaluation Education Programs**

Study 3 aimed to compile a list of formal postgraduate evaluation education programs (EEP) provided by academic institutions globally and to assess to what extent these programs cover the necessary content, topics, approaches, and methods to equip evaluation professionals with the skills needed to tackle challenges in an ever-changing world. Additionally, this study aimed to assess the extent to which the programs cater to the needs and interests of young and emerging evaluators (YEEs) regarding evaluation education. Specifically, the study examined the descriptions and curricula of university-based graduate evaluation education programs across various degrees such as diplomas, certificates, master’s, and doctoral degrees, as well as other operational characteristics such as location, modality, and disciplinary focus.

**Findings**

The study involved a comparison between the critical elements deemed vital by the interviewees and YEEs for a program aimed at equipping professionals to face the changing world and the features of the educational programs listed in the directory.

**Availability of Courses.** During the focus group, YEEs complained about the lack of availability of programs in their countries, especially in the Global South. To check this statement, the researcher classified the countries as part of the Global South and Global North, according to the Finance Center for South-South Cooperation list and counted the number of
courses per country. These data were organized in a separate database. The sample has 100 (42.55%) countries in the Global North and 135 (57.44%) countries in the Global South. Then, the researcher ran the chi-square goodness-to-fit test, where the expected proportion of courses is equal between the Global North and the Global South and found that the observed proportion is significantly different from the expected, \( \chi^2(1) = 4.33, p < .05 \). So, as the p value was less than .05, the difference between the two groups was considered significant. So, the results corroborate the YEEs’ claim.

**Thematic Focus.** Interdisciplinarity is crucial to equip evaluators with the knowledge to address the significant global challenges threatening the planet. Interdisciplinarity is the interaction among different bodies of knowledge or research practice (Committee on Facilitating Interdisciplinary Research and Committee on Science, 2005, as cited in Huutoniemi et al., 2010). The researcher adopted the assumption that programs that are generalist in nature are more likely to build an interdisciplinary vision among students. Therefore, evaluation education programs were classified into two types based on what they focused on: (1) generalist programs that give an overall education on evaluation without focusing on any specific field, even if they’re in a particular department or school; and (2) specialized graduate programs that concentrate on a narrow, specific area within a particular field, such as education, public health, or economics. The researcher employed a qualitative content analysis approach, which considered the names of colleges, schools, and departments where the programs were housed, as well as the programs’ descriptions, to classify them. Additionally, the researcher created a group of categories representing the disciplines of specialization. According to the results of the data classification, most of the programs were found to be specialized \( n = 207, 73.38\% \). Among the programs classified as "specialized," 43.00% \( n = 89 \) were housed in education-related disciplines,
14.01% \((n = 29)\) were housed in public administration-related disciplines, 13.53% \((n = 28)\) were housed in social sciences-related disciplines, 11.59% \((n = 24)\) were housed in health-related disciplines, and 9.18% \((n = 19)\) were housed in economics related disciplines. The remaining 8.70% \((n = 18)\) were dispersed across 14 other disciplines.

**Recommended Core Curriculum Dosage.** The purpose of analyzing evaluation education programs offered worldwide was to investigate the content taught and the differences among regions, types of courses, disciplines, and modalities. Due to the vast diversity in course descriptions and terminology used to name courses, answering this question presents a challenge. Interviewees share the opinion that formal education alone is insufficient for acquiring competence in the field. So, they emphasize the importance of exposure to a core curriculum. The essential components of a core curriculum include foundational knowledge, evaluation theories, basic statistics, and research methods. The analysis of the program descriptions and course names led to the rating of each core content element into three categories of inclusion: (1) explicitly declared; (2) suggested; and (3) not mentioned. First, each essential element was separately categorized according to the inclusion categories. In a subsequent step, the researcher defined five categories of dosage: (1) None of the four elements declared or suggested; (2) Elements other than foundational knowledge declared or suggested; (3) Only foundational knowledge declared or suggested; (4) Foundational knowledge and at least one other element declared or suggested; and (5) all four elements declared or suggested. Based on the results of the data analysis, the researcher concluded that a large proportion of the programs included in the study lack a complete core content recommended for evaluation education. Specifically, categories of dosage (1) and (2) demonstrate that despite having the term "evaluation" in their titles or courses, there may be a misalignment between the curricula and the consideration of
foundational principles of evaluation. However, this case is only evident in a small number of programs \( (n = 20, 7.38\%) \). Most programs \( (n = 193, 71.22\%) \) fall under categories of dosage (3) and (4), which suggests that they provide limited content considering the core curriculum. Only 48 \( (17.71\%) \) programs fall under category of dosage (5), offering an exemplary curriculum.

Further analysis was conducted to investigate potential relationships between the dosage level of core curriculum and type of program, region, disciplines, and modality, using various correlational analysis techniques. A Pearson’s chi-square test was performed to examine the relationship between core curriculum dosage level and type of program, assuming the hypothesis that there is no relationship between those two independent categorical variables. The test result showed a statistically significant association between the two variables, \( \chi^2 (12) = 42.04, p < .05 \). The likelihood ratio test produced a similar result, \( \chi^2 (12) = 43.73, p < .05 \). The Contingency Coefficient indicated a moderate strength of association, \( C = .37 \), while Cramer’s V suggested a small to moderate effect size, \( V = .23 \). These findings suggest a significant association between the two categorical variables. Table 6 presents the observed frequencies.

**Table 6**

*Core Curriculum Dosage Level Across Types of Programs (N=271)*

<table>
<thead>
<tr>
<th>Type of Program</th>
<th>Core content totally absent</th>
<th>Absence of Evaluation Foundations</th>
<th>Only evaluation foundation present</th>
<th>Foundational knowledge and at least one other element present</th>
<th>Complete core content offered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n )</td>
<td>%</td>
<td>( n )</td>
<td>%</td>
<td>( n )</td>
</tr>
<tr>
<td>1 Diploma</td>
<td>2</td>
<td>4.5%</td>
<td>2</td>
<td>4.5%</td>
<td>16</td>
</tr>
<tr>
<td>2 Certificate</td>
<td>0</td>
<td>0.0%</td>
<td>0</td>
<td>0.0%</td>
<td>9</td>
</tr>
<tr>
<td>3 Masters</td>
<td>2</td>
<td>1.6%</td>
<td>11</td>
<td>8.9%</td>
<td>10</td>
</tr>
<tr>
<td>4 Ph.D.</td>
<td>1</td>
<td>2.2%</td>
<td>2</td>
<td>4.4%</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>1.8%</td>
<td>15</td>
<td>5.5%</td>
<td>36</td>
</tr>
</tbody>
</table>
According to the data presented, masters programs have greater proportions of complete core content offered than all other types of programs.

A Pearson’s chi-square test was conducted to investigate the relationship between core curriculum dosage level and the region (Global South or Global North) where the programs are offered, assuming the hypothesis of no correlation. The test did yield a statistically significant result, $\chi^2 (4) = 16.46$, $p < .05$, indicating a significant association between the two variables. The likelihood ratio test produced a similar result, $\chi^2 (4) = 15.73$, $p < .05$. However, the contingency coefficient, $C = .24$, and Cramer’s $V$, $V = .25$, suggested a small to moderate effect size. These findings suggest a significant association between the two categorical variables. Table 7 summarizes the observed data.

**Table 7**

**Core Curriculum Dosage Across Regions (N=271)**

<table>
<thead>
<tr>
<th>Region</th>
<th>Categories of Analysis</th>
<th>1</th>
<th></th>
<th>2</th>
<th></th>
<th>3</th>
<th></th>
<th>4</th>
<th></th>
<th>5</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Core content totally absent</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Countries part of Global South</td>
<td>Absence of Evaluation Foundations</td>
<td>3</td>
<td>4.0</td>
<td>5</td>
<td>6.7</td>
<td>18</td>
<td>24.0</td>
<td>42</td>
<td>56.0</td>
<td>7</td>
<td>9.3</td>
</tr>
<tr>
<td>Countries not part of Global South</td>
<td>Only evaluation foundation present</td>
<td>2</td>
<td>1.0</td>
<td>10</td>
<td>5.1</td>
<td>18</td>
<td>9.2</td>
<td>125</td>
<td>63.8</td>
<td>41</td>
<td>20.9</td>
</tr>
<tr>
<td>Total</td>
<td>Complete core content offered</td>
<td>5</td>
<td>1.8</td>
<td>15</td>
<td>5.5</td>
<td>36</td>
<td>13.3</td>
<td>167</td>
<td>61.6</td>
<td>48</td>
<td>17.7</td>
</tr>
</tbody>
</table>

The group of countries part of Global North had greater proportions of complete core content offered than the group of countries part of Global South.

A Pearson’s chi-square test was performed to investigate the relationship between core content dosage level and disciplines to which programs are linked, assuming the hypothesis that there is no relationship between these two variables. The test result showed a statistically significant association between the variables, $\chi^2 (20) = 41.75$, $p < .05$, with a similar result, $\chi^2$
(20) = 44.724, $p < .05$ in the likelihood ratio test. The Contingency Coefficient indicated a moderate strength of association, $C = .36$, while Cramer’s $V$ suggested a small to moderate effect size, $V = .20$. These findings suggest a significant association between the two categorical variables. Table 8 summarizes the observed data.

**Table 8**

*Core Curriculum Dosage Level Across Disciplines (N=271)*

<table>
<thead>
<tr>
<th>Groups of disciplines</th>
<th>Categories of Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Core content totally absent</td>
</tr>
<tr>
<td>Education</td>
<td>0</td>
</tr>
<tr>
<td>Public Administration</td>
<td>1</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>0</td>
</tr>
<tr>
<td>Health</td>
<td>0</td>
</tr>
<tr>
<td>Economics</td>
<td>2</td>
</tr>
<tr>
<td>Others</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
</tr>
</tbody>
</table>

Proportionally, social science is the discipline that offers more complete core content compared to all the other disciplines.

A Chi-squared test of independence was conducted to analyze the relationship between modality and core content dosage. The analysis results indicated a significant association between modality and core content dosage $\chi^2 (12) = 29.07, p < .05$. The test also generated measures of effect size to understand the strength of the relationship between the variables. The Contingency Coefficient of .31 indicates a moderate association between modality and core content dosage, while Cramer’s $V$ value of .20 indicates a weak association between these variables. Table 9 summarizes the observed data.
These results suggest that the mode of delivery may influence the dosage of core content offered. While, advancements in technology have allowed for the development of online and blended learning formats that can also provide comprehensive core content in evaluation education, further studies are needed to deeper investigate why this is still not happening. Potential barriers might include limited technological infrastructure, lack of digital literacy and skills, pedagogical challenges, the need for practical skills and field experiences, accreditation and recognition concerns, faculty readiness and support, and cost considerations. Overcoming these barriers requires addressing issues such as infrastructure development, digital literacy promotion, faculty training and support, and financial investments. By doing so, institutions could harness the benefits of online and blended learning to expand access to comprehensive evaluation education.

Other aspects of a core curriculum that the interviewees mentioned and could have been analyzed were the alignment with the market, more local ownership of curricula, and less reliance on funding and donors. Regrettably, it was not feasible to analyze these two aspects.

### Table 9

<table>
<thead>
<tr>
<th>Modality</th>
<th>Categories of Analysis</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Core content totally absent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Absence of Evaluation Foundations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Only evaluation foundation present</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Foundational knowledge and at least one other element present</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Complete core content offered</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 On campus</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>2 Online</td>
<td>1</td>
<td>0.8</td>
<td>10</td>
<td>8.3</td>
<td>6</td>
<td>5.0</td>
</tr>
<tr>
<td>3 Mixed</td>
<td>2</td>
<td>3.2</td>
<td>10</td>
<td>8.3</td>
<td>13</td>
<td>27.6</td>
</tr>
<tr>
<td>4 Not Informed</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0.0</td>
<td>4</td>
<td>9.3</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>1.8</td>
<td>15</td>
<td>5.5</td>
<td>36</td>
<td>13.3</td>
</tr>
</tbody>
</table>

145
based on the data collected, as information on the development of the curricula, including the rationale and approach used, was not available in either the program descriptions or the course titles.

From the interviewees’ perspective, EEP should not have a standardized curriculum, leaving some degree of freedom for students to shape the program according to their interests and needs. Therefore, the programs were classified regarding the offering of elective courses by categorizing them as "yes" or "no." Assuming that a potential explanation for this difference was the type of program, a Pearson’s chi-square test was performed to examine the relationship between the type of program and elective classes offered. The test result indicated a statistically significant association between the variables, $\chi^2 (3) = 38.09$, $p < .05$. The likelihood ratio test produced a similar result, $\chi^2 (3) = 40.61$, $p < .05$. The Contingency Coefficient suggested a moderate strength of association, $C = .35$, while Cramer’s $V$ indicated a moderate effect size, $V = .38$. These findings suggest a significant association between the two categorical variables, and the data shows that doctorate degrees are the types of programs that proportionally offer more elective classes. Table 10 summarizes the observed data.

**Table 10**

*Programs Offering Elective Classes Across Types of Programs (N = 271)*

<table>
<thead>
<tr>
<th>Type of Program</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Elective classes offered</td>
<td>No elective classes offered</td>
</tr>
<tr>
<td></td>
<td>$n$</td>
<td>%</td>
</tr>
<tr>
<td>1 Diploma</td>
<td>7</td>
<td>15.9</td>
</tr>
<tr>
<td>2 Certificate</td>
<td>31</td>
<td>53.4</td>
</tr>
<tr>
<td>3 Masters</td>
<td>75</td>
<td>60.5</td>
</tr>
<tr>
<td>4 Doctorate</td>
<td>35</td>
<td>77.8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>148</td>
<td>54.6</td>
</tr>
</tbody>
</table>
**Topics Identified as Potentially Crucial for Evaluation Practice in the Future.**

Another aspect analyzed in the dataset was the presence of several topics identified by interviewees and YEEs as potentially crucial for evaluation practice in the future, including sustainable development and climate change, decolonization and indigenous systems of knowledge, systems thinking, conflicts and violence, social justice and human rights, workforce changes, and the effects of the COVID-19 pandemic. To assess the extent to which these topics were addressed in the sample of 271 programs under investigation, the researcher analyzed the program descriptions and the names of the required and elective courses, classifying them into three categories of inclusion: (1) explicitly declared, (2) suggested, and (3) not mentioned. The results, summarized in Table 10, indicate that a significant majority of programs (70.11%; \( n = 190 \)) did not declare or suggest any of the topics mentioned. These findings have important implications for the field of evaluation, as they suggest that there may be a need for increased attention to these emerging issues in evaluation education and training programs. Table 11 summarizes the observed data.
Table 11

*Programs that Declared or Suggested the Inclusion of Topics Identified as Potentially Crucial for Evaluation Practice in the Future (N = 271)*

<table>
<thead>
<tr>
<th>Topics</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>None of the topics cited</td>
<td>190</td>
<td>70.1</td>
</tr>
<tr>
<td>Social Justice and Human Rights</td>
<td>48</td>
<td>17.7</td>
</tr>
<tr>
<td>Systems Thinking</td>
<td>36</td>
<td>13.3</td>
</tr>
<tr>
<td>Sustainable Development and climate change</td>
<td>21</td>
<td>7.7</td>
</tr>
<tr>
<td>Decolonization</td>
<td>9</td>
<td>3.3</td>
</tr>
<tr>
<td>Conflicts and violence</td>
<td>5</td>
<td>1.8</td>
</tr>
<tr>
<td>Workforce</td>
<td>4</td>
<td>1.5</td>
</tr>
<tr>
<td>Covid-19</td>
<td>2</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Note: In this case, the total exceeds 271 because one program may include multiple topics.

In the second step, the researcher defined four categories of topics dosage: (1) None of the topics declared or suggested; (2) One topic declared or suggested; (3) Two topics declared or suggested; and (4) Three or more topics declared or suggested. Finally, the researcher investigated potential relationships between topics’ dosage level and type of program, region, and disciplines. A Chi-square test of independence was conducted to examine the relationship between topic dosage and type of program. The results showed a non-significant relationship, $\chi^2 (9) = 9.67, p = .378$ between the two variables. The likelihood ratio test also yielded non-significant results, $\chi^2 (9) = 11.39, p = .250$. The Contingency Coefficient was .19, and Cramer’s $V$ was .11, indicating respectively a weak association and effect size between the two variables. Therefore, the results suggest no statistically significant relationship between the type of program and dosage of topics variables.
A Pearson’s chi-square test was conducted to examine the relationship between the topic dosage and region where the programs are located. Results indicated a non-statistically significant association between the variables, $\chi^2 (3) = 6.50, p = .090$. The Contingency Coefficient indicated a weak association $C = 0.15$, and Cramer’s $V$ also showed a weak effect size ($V = .16$). The likelihood ratio test did not reach statistical significance ($\chi^2 (3) = 6.41, p = .090$). Therefore, the results suggest no statistically significant relationship between region where the program is offered (Global North and Global South) and dosage of topics variables.

A chi-square test of independence was conducted to examine the association between dosage of topics and disciplines. The results revealed a significant association between the variables, $\chi^2 (15) = 32.45, p < .05$. The Contingency Coefficient showed a moderate association ($C = .33$), while Cramer’s $V$ showed a weak effect size ($V = .20$). The likelihood ratio test also indicated a significant association between the variables, $\chi^2 (15) = 31.23, p < .05$. This result suggests that the two variables are related, and the relationship is not very strong. Table 12 summarizes the observed data.

**Table 12**

*Topic Dosage Level Across Disciplines (N = 271)*

<table>
<thead>
<tr>
<th>Groups of disciplines</th>
<th>Categories of Analysis</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>None of the topics</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>declared or suggested</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Education</td>
<td>72</td>
<td>75.0</td>
<td>20</td>
<td>20.8</td>
<td>4</td>
</tr>
<tr>
<td>Public Administration</td>
<td>24</td>
<td>68.6</td>
<td>4</td>
<td>11.4</td>
<td>4</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>31</td>
<td>70.5</td>
<td>8</td>
<td>18.2</td>
<td>5</td>
</tr>
<tr>
<td>Health</td>
<td>12</td>
<td>50.0</td>
<td>10</td>
<td>41.7</td>
<td>2</td>
</tr>
<tr>
<td>Economics</td>
<td>18</td>
<td>64.3</td>
<td>4</td>
<td>14.3</td>
<td>2</td>
</tr>
<tr>
<td>Others</td>
<td>33</td>
<td>75.0</td>
<td>6</td>
<td>13.6</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>190</td>
<td>70.1</td>
<td>52</td>
<td>19.2</td>
<td>20</td>
</tr>
</tbody>
</table>
Programs with three or more topics declared or suggested were relatively low across all disciplines, with the highest percentage found in Economics (14.3%), followed by Public Administration (8.6%), and Others (4.5%), and Education, Social Sciences and Health with no programs falling into this category.

**Approaches Likely to Gain Importance in the Field of Evaluation in a World in Transformation.** The interviewees mentioned seven types of approaches they believe will gain importance in the field of evaluation in a world in transformation: a systems-oriented approach; empowering, participatory, or deliberative approaches; developmental evaluation; contextually responsive evaluation; equity-focused evaluation; utilization-focused evaluation; and environmental approaches. To assess the extent to which these approaches were addressed in the sample of 271 programs under investigation, the researcher analyzed the program descriptions and the names of the required and elective courses, classifying them into three categories of inclusion: (1) explicitly declared, (2) suggested, and (3) not mentioned. In the second step, the researcher defined four categories of approaches dosage: (1) none of the seven approaches declared or suggested; (2) one of the seven approaches declared or suggested; (3) two of the seven approaches declared or suggested; and (4) three or more approaches of the seven declared or suggested. Finally, the researcher investigated potential relationships between approaches dosage level and type of program, region, and disciplines. Table 13 summarizes the overall observed data.
Table 13

Number of programs by Approach Level (N = 271)

<table>
<thead>
<tr>
<th>Approach level</th>
<th>Description</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>None of the Approaches suggested or declared</td>
<td>205</td>
<td>75.6%</td>
</tr>
<tr>
<td>2</td>
<td>One Approach suggested or declared</td>
<td>40</td>
<td>14.8%</td>
</tr>
<tr>
<td>3</td>
<td>Two Approaches suggested or declared</td>
<td>12</td>
<td>4.4%</td>
</tr>
<tr>
<td>4</td>
<td>Three or more Approaches suggested or declared</td>
<td>14</td>
<td>5.2%</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>271</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

The absence of specific evaluation approaches in the majority of program curricula (75.7%) can be attributed to various factors. One possible rationale is rooted in the dynamic nature of the evaluation field, characterized by the continual emergence of new approaches and methods. In response, the design of evaluation education programs may have prioritized imparting foundational principles and competencies that hold applicability across diverse approaches. This strategy avoids the endorsement of specific approaches that may become obsolete or less pertinent as the field progresses over time.

A more detailed descriptive analysis was conducted to see how many times each of the seven approaches were declared or suggested in the sample of 271 programs. Table 14 summarizes the observed data.
Finally, the researcher investigated potential relationships between approaches dosage and type of program, region, disciplines, and topics. A Chi-square test of independence was conducted to examine the relationship between approach dosage and type of program. The results showed that the chi-square statistic was non-significant, $\chi^2 (9) = 13.05, p = .161$, indicating no significant association between the two variables. The likelihood ratio test also yielded non-significant results ($X^2 (9) = 15.04, p = .090$). The Contingency Coefficient was .21, and Cramer’s $V$ was .13, indicating a weak association between the two variables. Therefore, the results suggest no statistically significant relationship between the approach dosage and type of program variables.

A Pearson’s chi-square test was conducted to examine the relationship between approach dosage and regions (Global South, Global North) where the programs are located. Results also indicated no statistically significant association between the variables, $\chi^2 (3) = .84, p = .839$. The contingency coefficient indicated a weak association ($C = 0.06$), while Cramer’s $V$ showed a weak effect size ($V = 0.06$). The likelihood ratio test did not also reach statistical significance ($\chi^2$...
(3) = .92, \( p = .821 \). Therefore, the results suggest that there is no statistically significant relationship between regions where the programs are located and approach dosage variables.

A chi-square test of independence was conducted to examine the association between approach dosage and disciplines. The results revealed no significant association between the variables, \( \chi^2 (15) = 12.69, p = .626 \). The likelihood ratio test also indicated no significant association between the variables, \( \chi^2 (15) = 15.14, p = .441 \). The Contingency Coefficient showed a weak association (\( C = .21 \)), while Cramer’s \( V \) showed a weak effect size (\( V = .13 \)). Overall, this suggests that the two variables are not associated.

A Pearson’s chi-square test of independence was conducted to investigate the relationship between topics and approaches dosage. The results showed significant results \( \chi^2 (9) = 215.40, p\text{-value} < .05 \). The likelihood ratio test indicated that the null hypothesis of independence between topics and approaches should be rejected, as the p-value was equal to 0. The Contingency Coefficient was calculated as .67, indicating a strong association between the variables. Cramer’s \( V \) coefficient was also calculated, and it was found to be .52, indicating a moderate effect size. Overall, the results suggest that there is a significant relationship between topics and approaches dosage. Table 15 summarizes the observed data.

**Table 15**

*Approaches Dosage Across Topics Dosage (N = 271)*

<table>
<thead>
<tr>
<th>Topics Dosage</th>
<th>Approaches Dosage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>None of the topics suggested or declared</td>
<td>174</td>
</tr>
<tr>
<td>One topic suggested or declared</td>
<td>25</td>
</tr>
<tr>
<td>2 topics suggested or declared</td>
<td>4</td>
</tr>
<tr>
<td>3 or more topics suggested or declared</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>205</td>
</tr>
</tbody>
</table>
From the observed results, the programs that presented low dosages of topics considered relevant where the ones that did not included relevant approaches. Based on the observed results, it was found that the programs with low coverage of topics considered relevant were those that did not include significant approaches.

**Research Methods.** The researcher examined program descriptions and course offerings to determine the focus of research methods taught by the programs. Like the previous analyses, the researcher employed a tripartite rating scheme, consisting of (1) explicit declaration, (2) suggestion, and (3) absence of mention, to classify the methods taught. The offer of quantitative, qualitative, and mixed methods in the EEP was analyzed and categorized according to the three inclusion categories. The findings are presented in Table 16, which reports the frequencies of methods declared or suggested in the analysis of course descriptions and courses offered by a sample of 271 programs. Of the 271 programs analyzed, 26.20% \((n = 71)\) did not declare or suggest any research method. The use of this classification scheme allowed for a thorough investigation of the programs’ research methodological foci.

**Table 16**

*Programs that Declared or Suggested the Inclusion of Specific Research Methods in the Curricula*

<table>
<thead>
<tr>
<th>Research Methods</th>
<th>(n)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantitative Methods</td>
<td>181</td>
<td>66.8</td>
</tr>
<tr>
<td>Qualitative Methods</td>
<td>145</td>
<td>53.5</td>
</tr>
<tr>
<td>Mixed Methods</td>
<td>36</td>
<td>13.3</td>
</tr>
<tr>
<td>No method</td>
<td>71</td>
<td>26.2</td>
</tr>
</tbody>
</table>

Note: In this case, the total exceeds 271 because one program may include multiple methods.

According to the interviewees, mixed methods are considered the most suitable approach for evaluations to contribute to transformative change. However, only 13.28% \((n = 36)\) of
programs declare or suggest that mixed methods are part of their curricula. This result is consistent with the trend toward quantitative methods in the evaluation field, as noted by the interviewees, who also highlighted the prevalent belief that only quantitative data could provide reliable evidence.

A Chi-square test of independence was conducted to examine the relationship between quantitative methods and disciplines. The results indicated that the observed frequency distribution did not significantly differ from the expected distribution, $\chi^2 (10) = 12.54, p = .251$. Additionally, the likelihood ratio and Pearson chi-square tests also revealed no significant association, $\chi^2 (10) = 12.67, p = .243$. Finally, the Contingency Coefficient was .21, indicating a moderate association level between the two variables, and Cramer’s V was .15, indicating a weak association level between the two variables. These results suggest that the two categorical variables under investigation are not significantly associated.

Another Chi-square test of independence was used to investigate the relationship between qualitative methods and disciplines. The results showed that the observed frequency distribution significantly differed from the expected distribution, $\chi^2 (10) = 23.36, p < .05$. The likelihood ratio tests also revealed a significant association, $\chi^2 (10) = 22.76, p = < .05$. The Contingency Coefficient was .28, and Cramer’s V was 0.21, indicating a moderate association. These results suggest a significant relationship between the two categorical variables being studied. The preference for qualitative research methods may vary across disciplines due to several factors. One potential reason is the epistemological orientation of different disciplines. Some disciplines, such as sociology, anthropology, and psychology, have traditionally emphasized understanding social phenomena in their natural contexts and exploring subjective experiences, making qualitative methods a better fit for their research objectives. Additionally, the nature of the
research questions asked within a discipline can influence the choice of research methods. Qualitative methods are often favored when studying complex social phenomena, exploring implementation processes, and gaining in-depth insights into individuals’ and communities’ perspectives. Conversely, disciplines that focus on quantifiable measurements, such as economics, may prioritize quantitative research methods for their ability to provide precise evidence and test hypotheses related to causal relationships between variables or comparisons among groups.

**New Technologies.** The interviewees and YEEs pointed out the use of new technologies as an important topic due to their capacity to save time and money, facilitate collaboration and knowledge exchange, promote lifelong learning, and surmount language barriers. In addition, these types of content are critical for evaluators to keep pace with the rapidly evolving evaluation landscape. To determine if evaluation education programs include the development of new technologies skills in their curricula, the researcher analyzed program descriptions and course offerings using three categories of inclusion: (1) explicitly declared; (2) suggested; and (3) not mentioned. The study found that only 53 programs (19.56%) declared or suggested the inclusion of new technologies, such as big data, artificial intelligence, machine learning, or data analysis applications. The researcher further explored the relationship between the inclusion of new technologies and region where the programs are located and type of program.

A Chi-square test of independence was conducted to analyze the relationship between the inclusion of new technologies and regions. The results indicated that the observed frequency distribution did significantly differ from the expected distribution, $\chi^2 (2) = 8.79, p < .05$. The likelihood ratio also revealed significant association, $\chi^2 (2) = 7.83, p < .05$. The **Contingency**
Coefficient was .18, and Cramer’s V was .16. So, these results suggest a significant association between the two categorical variables being studied. Table 17 shows the observed frequencies.

**Table 17**

*Inclusion of New Technologies in the Curricula Across Regions (N = 271)*

<table>
<thead>
<tr>
<th>Region</th>
<th>Categories of Analysis</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>New Technologies Explicitly declared as part of the curriculum</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Countries part of Global South</td>
<td></td>
<td>8</td>
<td>10.7</td>
<td>9</td>
</tr>
<tr>
<td>Countries not part of Global South</td>
<td></td>
<td>30</td>
<td>15.3</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>38</td>
<td>14.0</td>
<td>15</td>
</tr>
</tbody>
</table>

The data indicate that a higher proportion of programs situated in the Global South (22.7%) have declared or suggested the incorporation of new technologies within their curricula, comparing with programs situated in the Global North (18.4%).

Next, the researcher also conducted a Chi-square test of independence to explore the association between the inclusion of new technologies and the type of programs. The results indicated no significant difference between the observed and expected frequency distribution, $\chi^2 (6) = 11.33, p = .079$. However, the likelihood ratio test indicated a significant association between the two variables ($\chi^2 (6) = 16.65, p < .05$). In addition, the contingency coefficient was .2, and Cramer’s V was .15. These results suggest that there may be a weak relationship between the two categorical variables under investigation. Therefore, additional analysis is required to ascertain the nature and strength of this relationship. Due to the inconsistency observed between the Pearson and likelihood ratio tests, it is not feasible to make conclusive statements regarding the significance of the relationship.
In conclusion, while the importance of technology in evaluation has been stressed by participants of Studies 1 and 2, the study revealed that only a minority of programs explicitly include or suggest the incorporation of new technologies. This result highlights a potential gap in integrating technology in evaluation education that could hinder evaluators’ ability to adapt effectively to the changing landscape of their field. Future research could examine the reasons behind this gap and explore strategies for enhancing the integration of new technologies in evaluation education.

**Practicum and Soft Skills Development.** The opinion of most participants in this research, both interviewees and YEE participants in the focus group, was that practical application should be a mandatory requirement in evaluation courses and that through this practice, soft skills that are challenging to teach in theory would be acquired or at least experienced. For YEEs, a first experience would also be a gateway to the evaluation job market. However, due to their inexperience, they would require more than practice but also individual mentoring. Analyzing the program descriptions and courses offered allowed the researcher to identify if the programs included practicum in their curricula. The researcher used the same three rates categories of inclusion: (1) explicitly declared; (2) suggested; and (3) not mentioned. Overall, the practicum is declared as part of the curriculum for 112 programs (41.32%), and for 20 of them (7.38%), it is suggested. Additionally, only 14 programs (5.17%) declared mentorship as part of the curriculum.

The researcher investigated the relationship between the inclusion of practicum and regions and between the inclusion of practicum and the type of program. A Pearson’s chi-square test was conducted to examine the relationship between the inclusion of practicum and regions. The test result showed a statistically significant association between the two variables: $\chi^2 (2) =$
13.95, \( p < .05 \). The likelihood ratio test produced a similar result, \( \chi^2 (2) = 13.54, p < .05 \). The
Contingency Coefficient, \( C = 0.22 \) and Cramer’s V, \( V = 0.22 \), indicated a weak to moderate
strength of association. These findings suggest a significant association between the two
categorical variables but with a weak to moderate effect size. Table 18 summarizes the observed
data.

**Table 18**

*Practicum Offered Across Regions (N = 271)*

<table>
<thead>
<tr>
<th>Region</th>
<th>Categories of Analysis</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Practicum declared as</td>
<td>Practicum suggested</td>
<td>Practicum not</td>
<td></td>
</tr>
<tr>
<td></td>
<td>part of the curriculum</td>
<td>as part of the</td>
<td>mentioned</td>
<td></td>
</tr>
<tr>
<td>Countries part of Global South</td>
<td>n=20</td>
<td>n=11</td>
<td>n=44</td>
<td></td>
</tr>
<tr>
<td></td>
<td>26.7%</td>
<td>14.7%</td>
<td>58.7%</td>
<td></td>
</tr>
<tr>
<td>Countries not part of Global</td>
<td>n=92</td>
<td>n=9</td>
<td>n=95</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>46.9%</td>
<td>4.6%</td>
<td>48.5%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>112</td>
<td>20</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td></td>
<td>41.3%</td>
<td>7.4%</td>
<td>51.3%</td>
<td></td>
</tr>
</tbody>
</table>

The data show that a higher proportion of programs situated in the Global North (51.5%) have declared or suggested the incorporation of practicum within their curricula, comparing with programs situated in the Global South (41.4%).

The relationship between the inclusion of practicum and the type of program was examined using Pearson’s chi-square test. The results indicated a non-significant association between the two variables, \( \chi^2 (6) = 5.25, p = .512 \). The likelihood ratio test also revealed a non-significant result, \( \chi^2 (6) = 5.57, p = .473 \). The contingency coefficient was .14, and Cramer’s V was .10, indicating a small effect size.

While practicum content has the potential to foster the development of various skills, specific courses targeted at specific abilities have highlighted them more extensively. For example, a comprehensive analysis of 271 programs revealed that 54 (19.91%) declare they
include project management as a dedicated course, whereas 46 (16.97%) declare they incorporate communication-related content within their curriculum.

**Professionalization Aspects.** The insights that interviewees and YEE participants provided highlight the importance of incorporating professional aspects into evaluation education. This incorporation should cover a range of themes including ethics, access to communities of practice, access to referential documents such as guidelines, quality standards, and professional competencies, as well as opportunities for networking and employment. Next, the researcher examined the program’s description and courses to identify if the programs mentioned professionalization aspects. The researcher used the same three categories of inclusion: (1) explicitly declared; (2) suggested; and (3) not mentioned. Findings show that a small number of programs \(n = 58, 21.40\%\) declare or suggest that professional aspects are included in the content. Alignment with a set of competencies was declared or suggested by 11.07\% \(n = 46\) of the programs.

Lastly, the researcher studied the correlation between professionalization content inclusion and both regions and disciplines. The relationship between the inclusion of professionalization topics and regions was examined using a Pearson’s chi-square test. The results indicated a non-statistically significant association between professionalization content inclusion and regions, \(\chi^2 (2) = 2.02, p = .365\). The likelihood ratio test also revealed similar results, \(\chi^2 (2) = 1.93, p = .380\). The Contingency Coefficient was .09 and Cramer’s V was .09, indicating a weak effect size. Thus, it can be concluded that the inclusion of professionalization topics does not vary by region.

The relationship between the inclusion of professionalization topics and disciplines was examined using Pearson’s chi-square test. The results indicated a statistically significant
association between the two variables, $\chi^2(10) = 32.38, p < .05$. The likelihood ratio test also revealed similar results, $\chi^2(10) = 38.70, p < .05$. The contingency coefficient was .33, indicating a moderate effect size, and Cramer’s $V$ was .24, representing a small to moderate effect size. Thus, it can be concluded that the inclusion of professionalization topics varies by discipline. Table 19 summarizes the observed data.

**Table 19**

*Inclusion of Professionalization Across Disciplines (N = 271)*

<table>
<thead>
<tr>
<th>Groups of Disciplines</th>
<th>Categories of Analysis</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Professionalization topics declared</td>
<td>Professionalization topics suggested</td>
<td>Professionalization topics not mentioned</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Public Administration</td>
<td>2</td>
<td>5.7</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>5</td>
<td>11.4</td>
<td>7</td>
<td>15.9</td>
</tr>
<tr>
<td>Health</td>
<td>0</td>
<td>0.0</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td>Economics</td>
<td>4</td>
<td>14.3</td>
<td>7</td>
<td>25.0</td>
</tr>
<tr>
<td>Others</td>
<td>5</td>
<td>11.4</td>
<td>13</td>
<td>29.5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>17</strong></td>
<td><strong>6.3</strong></td>
<td><strong>41</strong></td>
<td><strong>15.1</strong></td>
</tr>
</tbody>
</table>

The data indicate that Economics (35.7%) and Social Sciences (29.5%) have the highest proportion of declared or suggested professionalization topics in their curricula compared to other disciplines. There are some possible reasons why certain disciplines may have a larger proportion of declared or suggested professionalization topics in evaluation education curricula compared to other disciplines. One reason could be the nature of the discipline itself, as some disciplines may inherently place a greater emphasis on practical applications and professional skills development. Furthermore, external factors such as societal demands and job market trends can influence the inclusion of professionalization topics in evaluation education curricula. As the
demand for skilled evaluators increases, disciplines may respond by incorporating more professionalization topics to meet the needs of employers and the evolving field of evaluation. It is important to note that these factors are context-dependent and may vary across disciplines and educational institutions. Further research and analysis are necessary to gain a comprehensive understanding of the specific reasons behind the variation in the proportion of professionalization topics among disciplines in evaluation education curricula. Table 20 shows a summary of the chi-square results.

Table 20

Summary of Chi-square Tests of Independence Results

<table>
<thead>
<tr>
<th>Core Curriculum Dosage</th>
<th>Types of Programs</th>
<th>Region (Global South/Global North)</th>
<th>Disciplines</th>
<th>Modalities</th>
<th>Elective classes offered</th>
<th>Topics Dosage</th>
<th>Approaches Dosage</th>
<th>Grant Methods</th>
<th>Qual Methods</th>
<th>New Technologies</th>
<th>Practicum</th>
<th>Professionalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>g²</td>
<td>df</td>
<td>p value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.373</td>
<td>-</td>
<td>&lt; .05 (*)</td>
<td>16.46</td>
<td>43.75</td>
<td>29.02</td>
<td>38.09</td>
<td>9.17</td>
<td>13.05</td>
<td>11.33</td>
<td>5.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dosage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g²</td>
<td>df</td>
<td>p value</td>
<td>3</td>
<td>9</td>
<td>9</td>
<td>0.378</td>
<td>0.161</td>
<td>0.079</td>
<td>0.512</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region (Global South/Global North)</td>
<td>g²</td>
<td>df</td>
<td>p value</td>
<td></td>
<td>6.50</td>
<td>0.84</td>
<td>8.79</td>
<td>13.95</td>
<td>2.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disciplines</td>
<td>g²</td>
<td>df</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modsities</td>
<td>g²</td>
<td>df</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elective classes offered</td>
<td>g²</td>
<td>df</td>
<td>p value</td>
<td></td>
<td>0.080</td>
<td>0.819</td>
<td>&lt; .05 (*)</td>
<td>&lt; .05 (*)</td>
<td>0.365</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topics Dosage</td>
<td>g²</td>
<td>df</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dosage</td>
<td>g²</td>
<td>df</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Conclusion

This study aims to investigate the extent to which evaluation education programs mapped worldwide cover the characteristics needed to equip evaluation professionals with the skills to tackle challenges in a world in transformation, as identified by participants in Study 1 and Study 2. The analysis begins by examining the availability of courses, revealing that only a limited number of countries offer specialized programs in evaluation. Furthermore, a comparison between the Global North and the Global South shows a significant disparity, with the Global
North offering a greater number of programs, corroborating the YEEs’ claim about the unavailability of courses in the Global South. The study then explores the thematic focus of evaluation education programs and finds that, although generalist programs are more likely to foster interdisciplinary vision among students, most programs tend to specialize in narrow, specific areas within particular fields, with education-related disciplines being the most common.

In terms of the core curriculum, which interviewees highlighted as crucial and inclusive of foundational knowledge, evaluation theories, basic statistics, and research methods, the data reveals that only a small number of programs offer a “complete” core curriculum. Further analysis explores potential relationships between core curriculum dosage level and program type, region (Global North or Global South), disciplines, and modality using correlational analysis techniques. The results show statistically significant associations between core curriculum dosage level and program types, with doctorate programs offering a greater proportion of complete core content compared to other types of programs. The analysis also reveals a significant association between core curriculum dosage level and region, with the Global North having the highest proportion of programs offering complete core content. Additionally, there is a significant association between core curriculum dosage level and disciplines, with social science disciplines offering more complete core content compared to other disciplines. The study investigates the inclusion of elective classes, which allow students to shape the program according to their interests and needs. The data shows that most programs offer elective classes, with a significant association between offering elective classes and program type, particularly with doctoral programs proportionally offering more elective classes.

Subsequently, the researcher conducted an analysis regarding the incorporation of emerging topics within the curricula of 271 evaluation education programs, including sustainable
development, climate change, decolonization, indigenous systems of knowledge, systems thinking, conflicts and violence, social justice and human rights, workforce changes, and the effects of the COVID-19 pandemic. The findings indicate that the majority of programs do not include or suggest the incorporation of these topics in their curricula. Furthermore, the analysis explores potential associations between the dosage of topics and program types, regions, and disciplines to which the programs are affiliated. The results obtained from chi-square tests reveal a lack of significant association between the dosage of topics and program types, as well as between the dosage of topics and regions. However, a significant association is observed between the dosage of topics and disciplines. Programs that declare or suggest three or more topics are relatively scarce across all disciplines. Among the disciplines, Economics exhibits the highest percentage, followed by Public Administration and Others. Notably, no programs in the fields of Education, Social Sciences, and Health fall into this category. These findings carry important implications for the field of evaluation, underscoring the need for heightened attention to these emerging themes in evaluation education programs.

In Study 3, an investigation was also conducted to examine the incorporation of seven distinct approaches to evaluation in the curricula of EEPs. These approaches were identified by interviewees and focus group participants as being crucial in a changing world. The data reveals that only a limited number of programs include at least one of the mentioned approaches. Furthermore, the study explored potential relationships between the dosage of approaches and variables such as program type, region, disciplines, and topics. Statistical analyses were performed, and the results indicate no significant associations between the dosage of approaches and program type, region, or discipline. However, a significant relationship was observed between topics and the dosage of approaches. The findings indicate that programs with low
coverage of relevant topics are the ones that lack significant incorporation of approaches. Overall, the data suggests that the analyzed programs have not extensively integrated the seven approaches to evaluation. The most frequently identified approaches among the included programs are Culturally Responsive Evaluation, Equity-focused Evaluation, and System-oriented Evaluation.

The study also conducted an examination of the research methods taught in evaluation education programs, classifying them into distinct categories. Among the 271 programs analyzed, a negligible proportion did not explicitly mention or suggest the inclusion of any specific research method in their curricula. Despite the interviewees considering mixed methods as suitable for facilitating transformative change in evaluations, only a limited number of courses declared or recommended the incorporation of such methods. Additionally, the study explored the relationship between disciplines and the utilization of quantitative and qualitative methods. The results revealed that, in the case of quantitative methods, the observed frequency distribution did not significantly deviate from the expected distribution. However, the statistical tests indicated a significant association between disciplines and the use of qualitative methods. Further investigation into the reasons why the preference for qualitative research methods may vary across disciplines presents an intriguing avenue for future research. One plausible explanation lies in the epistemological orientation of different disciplines. Disciplines such as sociology, anthropology, and psychology have traditionally emphasized comprehending social phenomena within their natural contexts and delving into subjective experiences, rendering qualitative methods more aligned with their research objectives. Another potential factor is the nature of the research questions posed within a discipline, which can influence the selection of
research methods. In both analyses, the strength of the observed relationships ranged from weak to moderate.

Furthermore, the study delved into the integration of new technologies in evaluation education programs. The findings revealed that only a minority of programs explicitly stated or suggested the incorporation of new technologies as part of their curricula. To explore the relationship between the inclusion of new technologies and the geographic regions where the programs are situated, tests of independence were conducted. The results demonstrated a significant deviation in the observed frequency distribution from the expected distribution. Specifically, a higher proportion of programs located in the Global South declared or suggested the integration of new technologies compared to programs in the Global North. However, the test of independence examining the association between the inclusion of new technologies and the type of programs yielded inconclusive results due to inconsistencies between the Pearson and likelihood ratio tests, necessitating further analysis. These findings underscore a potential disparity in the integration of technology within evaluation education, which could impede evaluators' capacity to adapt effectively to the evolving landscape of their field. Future research could explore the underlying reasons for this disparity and propose strategies to enhance the integration of technology in evaluation education.

Most participants in this research emphasized the importance of practical application and mentorship in evaluation education programs, highlighting their value in skill development and preparation for the job market, as indicated by the YEEs. Consequently, the study aimed to assess the extent to which practicum and mentorship are integrated into evaluation programs. The data revealed that nearly half of the programs explicitly included practicum as part of their curriculum, while only a small proportion incorporated mentorship. Additionally, the study
investigated the relationship between the inclusion of practicum and geographic regions, as well as the relationship between the inclusion of practicum and the type of program, using Pearson’s chi-square test. The results indicated a significant association between the inclusion of practicum and geographic regions, demonstrating a higher proportion of programs situated in the Global North that declared or suggested the incorporation of practicum within their curricula compared to programs located in the Global South. However, no significant association was found between the inclusion of practicum and the type of program. Furthermore, the study identified a limited number of courses that targeted specific abilities, such as project management and communication-related content.

This study also explored the importance of incorporating professionalization aspects into evaluation education and investigated the inclusion of such topics in evaluation education programs across regions and disciplines. The study found that a small portion of programs declare or suggest the inclusion of professionalization topics, and alignment with competencies is declared or suggested in even fewer programs. The researcher studied the correlation between the inclusion of professionalization content and regions and the inclusion of professionalization content and disciplines. The tests showed that the inclusion of professionalization topics does not varies by region. However, other tests indicated a statistically significant association between the inclusion of professionalization topics and disciplines. The results suggests that evaluation education programs should cover a range of professionalization themes, including ethics, communities of practice, and access to referential documents, and create opportunities for networking and employment.

The conclusion chapter of this dissertation brings together the findings from the three conducted studies, providing a comprehensive analysis of the research topic. It synthesizes the
insights and implications derived from empirical sources, contributing to the field's knowledge and offering valuable insights for future research and practical applications.
Chapter 5: Conclusion

This mixed-methods research has provided valuable insights into how formal evaluation education programs (EEP) respond to challenges in the evaluation field in a constantly transforming world. In addition, the findings of this research have been instrumental in analyzing how evaluation professionals have been educated to overcome contemporary challenges, such as addressing complexity, inequity, and sustainability, and producing high-quality evaluations that can drive meaningful change.

The Role of Evaluation in Contributing to Systemic Change and Implications for the Practice

The world’s transformations and their effects on our planet have been discussed in different areas and sectors in the last few years and various terms have been used to describe the concept of transformation, such as systemic or transformative change (Polanyi, 2001; Wright, 2004; McKibben, 2006; Kolbert, 2014). Participants were asked to explain their understanding of systemic or transformative change. Based on the perspectives shared by the interviewees, systemic change is a broad and interconnected change that alters existing paradigms, disrupts power relations, and brings attention to individuals or groups who were previously overlooked or marginalized. On the other hand, transformative change involves bringing excluded groups to the table, restructuring systems like racism and gender inequalities, and promoting transformative learning. The YEEs participating in the focus group stated that transformative change requires a paradigm shift from traditional linear models to systemic thinking, where all aspects are interrelated.
Evaluation, as the science of examining the value and worth of evaluands, holds the potential to contribute to a deeper understanding of systems, causality, and the formulation of effective solutions. The participants in this research acknowledged the role of evaluation in facilitating transformative change by providing critical information, evidence, and insights to comprehend complex issues, monitor progress, develop solutions, and refine strategies. However, for evaluations to fulfill this role, they must adopt a broader perspective that incorporates societal concerns like environmental sustainability and equity. It is essential for evaluations to be integrated into national evaluation systems embedded within governmental organizational structures to drive transformative change at a national level. Moreover, evaluators need to develop robust theories of change, be contextually responsive, and address equity and environmental sustainability considerations to bring about meaningful transformations. The interviewees and young and emerging evaluators emphasized the urgent need for a comprehensive and systemic evaluation approach. This approach entails moving away from linear and unidimensional models towards multidimensional theories of change. By embracing such an approach, evaluators can gain a better understanding of initiatives, their interconnectedness with other components of the system, and the multiple effects that different factors can produce.

Expanding upon the participants’ comprehension of transformation and their perspectives on the potential contribution of evaluation towards it, the interviewees were prompted to identify themes or topics that are anticipated to influence the trajectory of evaluation in the future. The topics include sustainable development and climate change, decolonization, conflicts and violence, social justice, human rights, workforce changes, and the impacts of the COVID-19 pandemic. As articulated by the interviewees, systemic change requires rethinking everyone’s
values and principles and adopting more ethical behaviors as a society. So, systemic change involves addressing critical issues such as the equitable distribution of natural resources, reduction of carbon emissions, incorporation of indigenous knowledge systems, and mitigation of displacement and the loss of lives resulting from wars, dictatorships, conflicts, and pandemics. Additionally, it requires the creation of a fairer distribution of wealth and equal opportunities for all members of society. The principles behind the topics listed are essential for realizing systemic change, and evaluation can be crucial in advancing this agenda.

Participants were asked to identify evaluation approaches that are expected to gain significance in the future. The specific approaches listed are closely aligned with the earlier mentioned topics. These approaches possess distinctive characteristics that make them particularly relevant for driving systemic change. Notably, they involve the integration of complexity and systems thinking into evaluation design, the incorporation of environmental considerations within evaluation processes, the inclusion of diverse stakeholders to amplify their voices and promote engagement, the adoption of adaptive and iterative approaches to accommodate dynamic environments, the responsiveness to contextual nuances, the prioritization of marginalized and underrepresented groups' needs, and the facilitation of actionable outcomes by promoting the utilization of evaluation findings. The connection between these approaches and the identified topics demonstrates their suitability for addressing the challenges and aspirations associated with transformative change.

Participants also discussed the methods likely to gain importance for evaluation practice in the coming years. The selection of data collection and analysis methods in evaluations are closely linked to the chosen approaches. In recent years, there has been a shift in perspective regarding evaluation methods, with increasing recognition that the exclusive use of quantitative
methods has significant limitations. The interviewees and YEEs recommend using qualitative or mixed methods to address quantitative methods’ limitations. Qualitative methods provide representation and a voice for different groups in evaluation processes, while mixed methods can provide more valid evidence, combining both qualitative and quantitative methods. For example, ground truthing, which involves verifying the validity of quantitative data interpretation, underscores the importance of direct observation and engagement in fieldwork to understand community experiences and needs comprehensively. The YEE participants also emphasized the cultural significance of storytelling approaches in the Global South and stressed the need to adapt evaluation methods to fit local cultural norms and values. Overall, the methods highlighted by the participants were aligned with the approaches listed before. Overall, evaluation processes must be related to evaluation purposes and principles to increase the likelihood of effective systemic change.

**Challenges that Lay Ahead for Evaluation as a Field in an Era of Transformational Change**

Up to this point, the questions discussed with participants focused on an ideal future scenario in which evaluation plays an essential role and can effectively contribute to transforming systems and improving societies. The literature review in this research mentions numerous theorists who have advocated that evaluation practice should be grounded in values of equity, sustainability, inclusion, and social justice. However, several challenges must be overcome to realize this scenario in the field. The challenges include transforming commissioners’ organizational cultures to embrace these values in their policies and strategies and elevating the evaluation field to a level of professionalization that is still evolving. Overall, evaluation must be established as a transdisciplinary field, and an evaluation ecosystem must be
built by articulating the necessary elements to make it a relevant practice to overcome the challenges. Advocating for the importance of evaluations to bring about changes, addressing political aspects, building capacity correctly, and positioning evaluation in strategic roles in organizations and national governments are also critical steps toward achieving this goal.

**Implications for Evaluation Education**

Multiple avenues exist to achieve the envisioned ideal scenario of evaluation practice playing a pivotal role in transforming systems and improving societies. Evaluation education is widely recognized as a critical element in preparing professionals capable of conducting high-quality evaluations. This recognition justifies why evaluation education is one of the six critical components for professionalization, as identified by the United Nations Evaluation Group (UNEG, 2016; Picciotto, 2011). However, the task of education and training is challenging, given that the evaluation field consists of professionals with diverse backgrounds who entered the field through various pathways. Despite previous discussions in the literature, there is no globally recognized set of competencies to define a program that can adequately prepare professionals to address the complex challenges of evaluations. Thus, questions arise regarding what constitutes an exemplary evaluation education program, including the requisite curriculum contents, whether limited to evaluation fundamentals or inclusive of other complementary topics, and the opportunity to apply theoretical knowledge in practice.

The interviewees and YEEs presented a range of perspectives on the potential design of evaluation education programs aimed at equipping professionals with the requisite skills to navigate evaluations amidst a rapidly evolving world. From their perspective, a complete core curriculum comprises foundational evaluation knowledge, evaluation theories, basic statistics, and research methods. Evaluation education curricula should also be based on competencies
when available in each country, or ideally, a set of internationally agreed-upon competencies. However, interviewees also state that a core curriculum is insufficient to prepare professionals for the rapidly changing evaluation landscape. It is also necessary to have an interdisciplinary background, an understanding of systems change, and contextual responsiveness. Consequently, evaluation education programs must include a broad range of topics, theories, methods, and tools to effectively equip professionals with the requisite skills to navigate evaluations amidst a rapidly evolving world. Evaluators must also learn to select appropriate approaches based on context, feasibility, and stakeholders’ interests and use mixed methods, big data, technological applications for data analysis, and data visualization techniques to navigate evaluations in a rapidly evolving world.

Furthermore, theoretical knowledge is insufficient to equip professionals with the requisite skills to navigate evaluations amidst a rapidly evolving world. Evaluators must learn to apply theoretical knowledge to real-world evaluation scenarios, learn from mistakes, and develop critical thinking abilities. As an applied science, evaluation requires skills primarily related to interpersonal relations, process management, and communication. Practitioners and academic experts believe these skills can be acquired through practical exposure to programs and project evaluation, role-play, and on-the-job training. By gaining these skills, evaluators can be better equipped to design and implement evaluations that are culturally sensitive, inclusive, and responsive to the needs of diverse stakeholders, making a meaningful contribution to systems change. So, incorporating practicum in evaluation education programs is crucial for soft skills development. Therefore, another recommendation was that the knowledge acquired through practicum be part of the assessment systems in evaluation education programs.
Discussions about teaching modality concluded that the most recommended modality of evaluation education to fulfill potential students’ needs, especially YEEs’, is hybrid — a mix of face-to-face and online experiences. Face-to-face modalities provide opportunities for interactive discussions, interpersonal skills development, and observing experienced evaluators in action. Online modalities offer access to knowledge, accommodations of personal contexts, connections with diverse instructors and classmates, and the development of digital literacy skills.

During the discussions, other complementary aspects emerged, leading to several recommendations. Firstly, evaluation education programs should acknowledge the varying training needs of different professional roles in the field. For instance, external evaluators, scholars, and commissioners should receive specific training. Secondly, evaluation education programs should incorporate topics related to evaluation professionalization, including the benefits of joining VOPEs, which focus on initiatives related to sharing practices, developing standards, competencies, and ethical guidelines in evaluation. Additionally, YEEs emphasized the need for educational programs to include career support for students by guiding them on how to apply for evaluation positions or respond to Requests for Proposals, better preparing them for the labor market.

Therefore, to perform evaluations that foster transformative change, professionals must possess contemporary knowledge and skills, be able to self-assess their competencies, and evaluate the adequacy of formal evaluation education programs. Furthermore, previous experience and maturity are also crucial elements.

The expectations of YEEs regarding evaluation education programs revolve around practical experiences and opportunities to apply evaluation techniques in real-world situations, along with access to tools for all stages of evaluations. YEEs highly value practical experience,
which facilitates the development of professional networks, exposure to diverse evaluation contexts, and career advancement opportunities. In addition, exposure to diverse evaluation contexts broadens YEEs’ perspectives and cultural competence, which is critical in today’s globalized and diverse evaluation environment. YEEs also desire the integration of evaluation education into higher education and research methodologies to better understand the frontiers between research and evaluation. Moreover, YEEs seek inclusive and financially accessible programs.

Study 3 utilized the insights provided by participants regarding the design of educational programs aimed at equipping professionals with the necessary skills to navigate evaluations in an ever-changing world. The study employed a sample of 271 programs from across the world, analyzed based on the criteria extracted from the interviews and focus group data. The analysis commenced by examining the availability of programs, showing that only a small number of countries offer graduate programs specializing in evaluation, and Global North is the region that offers proportionally a larger number of programs. Regarding programs’ foci, participants emphasized the importance of interdisciplinary knowledge to equip evaluators with the tools to address the significant global challenges threatening the planet. Based on a content analysis, most of the programs are specialized with a considerable proportion of these programs housed in education-related disciplines. Other disciplines with specialized programs included public administration-related disciplines; social sciences-related disciplines; health-related disciplines; and economics-related disciplines.

The third aspect of the analysis was focused on identifying the four essential components that should be included in the core curriculum of an evaluation education program. These components are foundational knowledge, evaluation theories, basic statistics, and research
methods. The results show only a tiny subset offering a comprehensive core curriculum.

Additionally, the study investigated the provision of elective courses within evaluation education programs. As per the interviewees’ recommendations, evaluation education programs should give students the flexibility to tailor their curriculum to meet their interests and requirements. The findings reveal significant variability in the elective courses offered by evaluation education programs, with more than half of the programs offering this type of courses.

The fourth aspect analyzed in this study was the presence of emerging topics in evaluation education programs, including sustainable development and climate change, decolonization and indigenous systems of knowledge, systems thinking, conflicts and violence, social justice and human rights, workforce changes, and the effects of the COVID-19 pandemic. The results indicate that a significant proportion of programs did not mention or suggest incorporating any of these topics, highlighting a need for increased attention to these emerging issues in evaluation education programs. A fourth aspect examined the presence of seven different approaches to evaluation deemed essential in a changing world by interviewees. Again, the findings reveal that most programs did not declare or suggest including any of these approaches. A strong correlation was found between the lack of approaches included and the lack of attention to emerging topics.

Regarding research methods taught in evaluation education programs, although mixed methods were considered the most suitable for evaluations to contribute to transformative change by interviewees, only a tiny proportion of courses declared or suggested their inclusion in the curriculum. The results also indicate that the trend toward quantitative methods in the evaluation field is reflected in the descriptive statistics.
Despite the emphasis placed by many interviewees on the importance of technology for contemporary evaluation practice, the study found that only a minority of programs explicitly include or suggest the incorporation of new technologies. This result highlights a potential gap in integrating technology in evaluation education that could hinder evaluators’ ability to adapt effectively to the changing landscape of their field. Future research could explore the reasons behind the lack of new technologies as part of EEP and examine strategies for enhancing its integration in evaluation education.

According to the interviewees and the YEEs, practical application and mentorship are essential components of evaluation education programs for developing skills and preparing students for the job market. However, the data from the study indicate that while a moderate part of the programs declared the inclusion of practicum in the curriculum, only a small number explicitly mentioned mentorship.

Additionally, the study examined the incorporation of professionalization themes into evaluation education programs, such as ethical behavior, referential documents for practice (e.g., competencies, guidelines, and standards), and the importance of belonging to a community of practice. The results suggest that only a small number of the programs declared or suggested including these themes, indicating a lack of coordination between Voluntary Organizations for Professional Evaluation (VOPEs) and universities.

In conclusion, the findings of this study are significant for fostering a critical perspective on the congruity between current trends and the content of programs designed to prepare professionals for the evolving evaluation landscape worldwide. Furthermore, these results prompt further discussions on modifying the practice and teaching of evaluation. Specifically, the study reveals substantial deficiencies in including pertinent topics, approaches, and methods
in evaluation education programs globally and a weak connection with professionalization. Therefore, it is recommended that evaluation education programs adopt a more comprehensive approach by integrating emerging issues, new technologies, practicum, mentorship, and professionalization themes to better equip evaluation professionals for the contemporary demands of the field.
Chapter 6: Discussion

The provision of pre-service evaluation education is undeniably essential for preparing competent professionals who can produce high-quality evaluations across various disciplines and meet clients’ expectations (Patton, 1990; Gullickson et al., 2019). Moreover, such education plays a crucial role in equipping evaluation professionals with the necessary skills to address the challenges posed by a rapidly transforming world. Despite the expansion of evaluation education programs worldwide, they still exhibit several gaps and limitations (LaVelle, 2014; Meyer, 2016; Gullickson et al., 2019; King and Ayoo, 2020). For instance, technical skills like data collection and statistics are emphasized more than soft skills (e.g., project management, ethics, communication, conflict resolution, and multitasking). Furthermore, many programs lack supervised internships, which are essential for learners to acquire practical experience and develop soft skills. Additionally, the field of evaluation education lacks a consensus on what constitutes an adequate evaluation training program (King and Ayoo, 2020). Challenges include the absence of a universally accepted set of competencies for evaluation professionals and the lack of accreditation processes that connect standards, competencies, and curriculum in educational institutions (Schwandt, 2015).

The Anthropocene era has brought about profound changes in social relations and economic structures, with excessive resource exploitation leading to long-term environmental impacts such as climate change, deforestation, and pollution (Polanyi, 2001; McKibben, 2006). International organizations have recognized the risks these transformations pose to life on Earth and have therefore initiated pacts and treaties aimed at addressing global issues, mobilizing societies across countries to contribute to these efforts. At the national level, these organizations have called upon governments to establish monitoring and evaluation systems to track the
implementation of these treaties, emphasizing the importance of inclusive, transparent, and participatory follow-up mechanisms (Meyer, 2020). Additionally, large-scale initiatives to transform global systems and address complex global challenges have been primarily implemented by international donors and development agencies, leading to a growing demand for evaluations in this domain. It has become evident that the field lacks professionals with the appropriate competencies to evaluate these complex endeavors. The need to incorporate complexity and systems thinking (Bamberger et al., 2016), embrace new understandings, approaches, and data collection methods (Mertens, 2008; Chilisa, 2015; Meyer, 2020; Magro et al., 2021), and be responsive to local contexts has emerged as transformative elements shaping the practice and conception of evaluation.

The findings from the interviews and focus group discussions conducted in this study align with the insights derived from the literature review, emphasizing the implications of global transformations on evaluation practice and evaluators, as well as the competencies that need to be reinforced in addition to those already developed in evaluation courses. This alignment reinforces the significance and relevance of the present study.

**Strengths**

This research has three main strengths. First, as the study focuses on systems change, participants were prompted to consider future perspectives. This approach allowed the analysis to center on where we are now and the necessary changes to achieve a desired future scenario. Second, the study employed qualitative research to examine evaluation education programs. Semi-structured interviews with 19 specialists from diverse continents, backgrounds, and involvement in VOPEs provided insights into the current challenges of the evaluation field, as well as changes necessary in evaluation practice and education to contribute to systems change.
Third, the focus group with 11 YEEs from different continents provided a valuable perspective on their expectations, challenges, and barriers to accessing evaluation education. Giving voice to YEEs was crucial in this research because they represent the future workforce of the evaluation field. Finally, both the interviews and focus group discussions addressed the potential role of evaluation in contributing to systems change and the implications for evaluation practice and education. Overall, the qualitative data generated functional criteria to assess the extent to which existing evaluation education programs address relevant topics, approaches, and methods to prepare evaluation professionals for the field’s challenges and whether the current landscape of evaluation education meets YEEs’ expectations and needs. No previous studies have provided such a comprehensive list of criteria for evaluating evaluation education.

Another key strength of this research was the creation of an updated global directory of academic postgraduate evaluation programs across diverse countries, modalities, disciplines, and languages. This directory, publicly accessible on the Global Evaluation Initiative (GEI)’s Better Evaluation knowledge platform, is intended to offer easily accessible information on academic evaluation programs to a variety of stakeholders, including prospective students seeking to enhance their evaluation competencies and skills, public and private (for-profit and non-profit) staff seeking to develop their evaluation capabilities, evaluation education program administrators and academic staff interested in peer learning and knowledge exchange, representatives from VOPEs hoping to disseminate information on evaluation learning opportunities, evaluation scholars and other researchers interested in understanding the scope and diversity of academic training opportunities in evaluation, and evaluation champions and other

8 https://www.betterevaluation.org/academic-programs-directory
stakeholders committed to advancing the field and practice of evaluation. Therefore, this database offers a valuable and accessible product to the global evaluation community.

**Limitations**

Despite the favorable aspects of the study presented, significant limitations are evident. Firstly, the purposive sampling strategy employed in Study 1 may not be entirely representative of the views held by the global evaluation community across diverse regions and professional roles. Furthermore, utilizing a snowball sampling approach may have resulted in selection bias in the choice of interviewees. Similarly, recruiting YEEs participants for the focus group through social media advertising and mailing lists may have excluded individuals from vulnerable populations, diverse linguistic backgrounds, and various geographical locations. Another limitation pertains to the database of EEP, as the researcher relied on previous directories cited in academic literature, potentially omitting any programs created after their publication dates. Additionally, relying on lists disseminated by VOPEs as another source of data may imply using outdated lists. Moreover, the data collection process depended on information provided by official institutions’ websites, which may be incomplete, outdated, or biased due to marketing purposes. These inaccuracies may have influenced the content analysis.

**Implications and Future Directions**

This research provides some insight to the challenge of distinguishing EEP that comprehensively prepare evaluation professionals from those that merely reinforce specialization and limit critical thinking. The study’s outcomes should encourage EEP formulators and administrators to reflect critically on the suitability of their curricula and graduates’ preparedness to perform different roles in the market. They should also consider whether their students are ready to face future challenges in the field. Students’ readiness requires the ability to engage
stakeholders, adopt culturally and equity-responsive approaches, utilize mixed methods, and master new technologies toward a more just and sustainable world. Additionally, this research highlights the need to connect evaluation education with the field’s professionalization. For instance, developing a set of globally agreed-upon competencies that can be adapted and supplemented in national contexts would be beneficial for designing and evaluating EEP. Therefore, a future direction for the evaluation community is to continue working together to advance this discussion.

This research has important implications for the selection process of potential students who seek to balance contextual factors with their aspirations for evaluation education. Specifically, two distinct groups of individuals should be considered. Younger and less experienced individuals may need to maintain employment while acquiring experience to advance their careers in the field, which poses challenges in reconciling their goals with the demands of postgraduate education. Conversely, more experienced and established professionals may desire to supplement their existing competencies with contemporary and advanced content, potentially dedicating personal resources to further their education. These two groups of individuals are likely to employ different criteria in selecting postgraduate programs, highlighting the importance of program formulators hearing the voices of different target populations to calibrate education strategies continuously.

This research offers several potential directions for future research on evaluation education. First, a more in-depth analysis of the content, methods, and rationale of evaluation education is necessary to shed light on future directions. It is crucial to understand the reasons for different curricular conceptions among disciplinary approaches, types of programs, languages, and cultural factors, to assess the adequacy of evaluation education programs across
different contexts. Some critical research questions that could be explored include why some evaluation education programs lack foundational evaluation knowledge, why there is a significant difference in access to programs between the Global North and South, and how to address these issues. Additionally, further investigation into academic institutional challenges in promoting interdisciplinarity and transdisciplinarity, practicum and mentorship opportunities, and faculty member profiles is needed. Finally, exploring the cultural characteristics that differentiate evaluation practice and education across different regions and continents could be a fruitful area of investigation because it can reveal the underlying values, beliefs, and norms that shape these practices and help us develop more culturally responsive and effective evaluation education strategies.
References


https://doi.org/10.1016/j.evalprogplan.2018.10.009


https://doi.org/10.1057/9781137376374_8


https://doi.org/10.1002/ev.20466


http://catdir.loc.gov/catdir/enhancements/fy0658/89010426-t.html


https://doi.org/10.1016/j.evalprogplan.2019.02.012


https://doi.org/10.1177/1356389011403362


http://site.ebrary.com/id/10375064


Rodríguez-Bilella, P., Mulder, S.S., Zaveri, S. (2021). To Be or Not to Be an Evaluator for


[https://doi.org/10.4102/aej.v8i1.476](https://doi.org/10.4102/aej.v8i1.476)

Appendix A

Interview Protocol
Qualitative Study of Formal Evaluation Education Programs

**Introduction and Consent:** Remind that we will be recording the interview to refer to during the study but will delete the recording by June 2022. If any specific details need to be obfuscated, please let me know in the interview, and we will ensure that only the obfuscated results are included in transcripts and analysis.

---

**Background:**

Name: (first name only as an ID)
Organization:
Current Title:

**Section 1**

**Goal: 5 minutes warming up and talking about the professional background**

1. Tell me about your professional background in evaluation?
2. What do you consider your main contribution to the field of evaluation?

   Possible prompts:
   
   - Leader of a regional or national VOPE
   - Scholar (university faculty member and researcher)
   - Evaluation practitioner
   - Commissioner
   - Relevant published works

**Section 2**

**Goal: 10 minutes general discussion about the interviewees’ perspective on the role of evaluation**

3. In the last years, there have been discussions among the evaluation community about the role evaluation can play in supporting changes on systems related to critical issues of the planet, such as sustainability, inequality, social justice, etc. What is your perspective regarding this idea?

   Possible prompts: Prague Declaration signed in 2019 during IDEAS conference, Blue Marble Evaluation, other VOPEs conferences.
Section 3

Goal: 15 minutes general discussion about the implications for evaluation practice

If the interviewer agrees with the idea:

4. Given the call for evaluation to contribute to systemic changes, what are the main challenges for the evaluation practice?

5. Given this role, which themes, methods, and approaches are likely to gain more importance for the evaluation practice in the coming years?

If the interviewer disagrees with the idea:

5. Which themes, methods, and approaches are likely to gain more importance for the evaluation practice in the coming years?

Possible prompts:

- Themes: sustainable development (Agenda 2030), climate change; equity; social justice; complexity; systems thinking; sustainability; power differences; ethics; communication; social network analysis; stakeholders’ facilitation
- Approaches: culturally responsive evaluation; equity focused evaluation; gender responsive evaluation; indigenous evaluation; participatory evaluation; transformative evaluation; evaluation capacity building; developmental evaluation; blue marble evaluation
- Methods: mixed-methods, qualitative, quantitative

6. What formal evaluation education programs’ curricula should include to preparing better these professionals (it includes formal evaluation education programs at university level master’s, doctorate, certificates, and diplomas)?

   a. What essential themes do you think should be included? For example, do you think evaluation should expand to areas other than social programs?
   b. What essential approaches do you think should be taught?
   c. What methods do you think should be taught?
   d. What is your opinion about including practicum (should it be required)?
   e. What is your opinion about the teaching method (online, on-campus, hybrid/blended)

Goal: 5 minutes of wrap-up

7. Is there anything else you would like to add about evaluation education?

8. Please nominate other three to four international experts to be interviewed in this study who play a relevant role in the field representing the following groups:
   i. the leadership of formal evaluation education programs.
ii. Research on evaluation.

iii. Authorship of relevant books and articles that discuss evaluation as a professional field

iv. Leaders of organizations for professional evaluation (VOPEs); and

v. Multicultural (international) evaluation practitioners and commissioners.

Closing

I am incredibly grateful for the time you have taken with me today. I hope this study will shed light on the discussions about how formal evaluation education programs respond to new challenges in the field of evaluation in a context of a world in transformation, and your contribution today was invaluable.
Appendix B

Draft Email Invitation for Interview – First Subjects

Subject: Interview request regarding formal evaluation education programs

Dear [Expert],

As a Ph.D. student at Claremont Graduate University, I am working on a research project exploring how formal evaluation education programs respond to the challenges in the field of evaluation in a context of a world in transformation. This study will contribute to a comprehensive reflection on formal evaluation education programs.

You have been selected for your contribution to the field of evaluation in recent years as a leader of an International Organization with the mission to advance evaluation as a profession across the globe. I would like to know if you are available to share your insights through an interview in the next few days. If you are willing to participate, please let me know your availability. You will be asked about your perspectives on the role of evaluation in supporting systems changes and the implications for evaluation practice and education.

The interview will be held in English. According to your preference, it will be conducted via phone or video call application (zoom, google meets, or teams, at your convenience) and should take less than an hour. I will be voice-recording the interview to ensure that my notes are accurate, but recordings will be deleted after the transcription (by June 2022).

At the end of the interview, you will be asked to nominate three to four other experienced evaluation experts who have contributed to the debate on the importance and role of evaluation in improving initiatives, policies, and change systems over the last years. The nominated experts will represent the following groups: leadership of formal evaluation education programs; researchers on evaluation; authors of books and articles that discuss evaluation as a professional field; leaders or formal leaders of voluntary organizations for professional evaluation (VOPEs); and multicultural evaluation practitioners and commissioners. The Informed Consent form is attached for your review and gives additional information on the study.

Thank you very much for your time.

Sincerely,

Marcia P Joppert, MPA
marcia.joppert@cgu.edu
Ph.D. Student in Evaluation and Applied Methods
Claremont Graduate University
+1(909) 997-0235
Appendix C

Draft Email Invitation for Interview – Referred Subjects

Subject: Interview request regarding formal evaluation education programs

Dear [Expert],

As a Ph.D. student at Claremont Graduate University, I am working on a research project exploring how formal evaluation education programs respond to the challenges in the field of evaluation in a context of a world in transformation. This study will contribute to a comprehensive reflection on formal evaluation education programs.

You were referred to as someone who has contributed to the debate on the importance and role of evaluation in improving initiatives, policies, and change systems over the last years, representing the following groups: leadership of formal evaluation education programs; researchers on evaluation; authors of books and articles that discuss evaluation as a professional field; leaders or formal leaders of voluntary organizations for professional evaluation (VOPEs); and multicultural evaluation practitioners and commissioners.

I would like to know if you are available to share your insights through an interview in the next few days. If you are willing to participate, please let me know your availability. You will be asked about your perspectives on the role of evaluation in supporting systems changes and the implications for evaluation practice and education.

The interview will be held in English. According to your preference, it will be conducted via phone or video call application (Zoom, Google meets, or Teams, at your convenience) and should take less than an hour. I will be voice-recording the interview to ensure that my notes are accurate, but recordings will be deleted after the transcription (by June 2022).

The Informed Consent form is attached for your review and gives additional information on the study.

Thank you very much for your time.

Sincerely,

Marcia P Joppert, MPA
marcia.joppert@cgu.edu
Ph.D. Student in Evaluation and Applied Methods
Claremont Graduate University
+1(909) 997-0235
Appendix D

Draft Email Reminder for Interview

Subject: Interview Reminder

Dear [expert],

You are scheduled for a Phone/Zoom/Google Meets/Teams (Select one) meeting on XXXX at XXX. Please let me know if you are still available at this time.

The interview should take approximately 45 minutes and be conducted via Phone/Zoom/Google Meets/Teams (Select one). The interview will be recorded, with your permission, to ensure that my notes are accurate, but recordings will be deleted at the end of this project (by June 2022).

Our goal is to explore how formal evaluation education programs respond to the challenges in the field of evaluation in a context of a world in transformation. This study is qualitative, and our research questions are:

a. Given the call for evaluation to contribute to systemic changes, what are the main challenges for the evaluation practice?

b. Given the call for evaluation to contribute to systemic changes, what are the topics, approaches, and methods that should be taught to prepare professionals to respond to the field’s challenges?

c. What are the needs and interests of the YEEs concerning the most relevant topics, approaches, and methods pointed out by international evaluation experts?

The Informed Consent form is attached for your review and gives additional information on the study.

Thank you for your time.

Sincerely,

Marcia Joppert, MPA
marcia.joppert@cgu.edu
Ph.D. Student in Evaluation and Applied Methods
Claremont Graduate University
+1(909) 997-0235
Appendix E

Draft Interview Consent Form

[Claremont Graduate University logo]

AGREEMENT TO PARTICIPATE IN A STUDY OF FORMAL EVALUATION EDUCATION PROGRAMS (IRB # 4164)

You are invited to participate in a research project. Volunteering will not benefit you directly, but you will be helping the investigators to understand better how formal evaluation education programs are responding to the challenges in the field of evaluation in a context of a world in transformation. If you decide to volunteer, you will be asked to participate in an interview about your perspectives on the role of evaluation in supporting systems change and the implications of this idea for the evaluation practice. The interview will take about 45 minutes of your time. Volunteering for this study involves no more risk than what a typical person experiences on a regular day. Your involvement is entirely up to you. You may withdraw at any time for any reason. Please continue reading for more information about the study.

STUDY LEADERSHIP: This research project is led by Marcia Joppert of the Claremont Graduate University, supervised by professor Dr. Stewart Donaldson.

PURPOSE: This study aims to understand better how formal evaluation education programs respond to the challenges in the field of evaluation in a context of a world in transformation.

ELIGIBILITY: To participate in this study, you should have contributed to the field of evaluation in recent years as a scholar, practitioner, commissioner, and leader of a Voluntary Organization for Professional Evaluation (VOPE) in your country or region or internationally.

PARTICIPATION: During the study, you will be asked to interview about your perspectives on the role of evaluation in supporting systems changes and the implications for evaluation practice and education. The interview will take about 45 minutes of your time.

RISKS OF PARTICIPATION: The risks you run by taking part in this study are minimal. The interview topics ask you to discuss your perspectives on the field of evaluation. Your name will not be associated with these perspectives in the study.

BENEFITS OF PARTICIPATION: We do not expect the study to benefit you personally. This study will benefit the researcher by fulfilling her academic requirements. This study also intends to benefit the field of evaluation by contributing to the knowledge base of research on evaluation.

COMPENSATION: You will not be directly compensated for participating in this study.

VOLUNTARY PARTICIPATION: Your participation in this study is entirely voluntary. You may stop or withdraw from the study or refuse to answer any question for any reason at any time without it being held against you. Your decision whether to participate or not will have no effect on your current or future connection with anyone at CGU or with any other relevant entity/agency.

CONFIDENTIALITY: Your individual privacy will be protected in all papers, books, talks, posts, or stories resulting from this study. We may use the data we collect for future research or share it with other researchers, but we will not reveal your identity. In addition, the voice-recording of the interviews will be used for transcription purposes and will be erased after it to protect your privacy.
SPONSORSHIP: This study is not being sponsored.

FURTHER INFORMATION: If you have any questions or want additional information about this study, please contact Marcia Joppert at marcia.joppert@cgu.edu, +1(909) 997-0235. You may also contact Dr. Stewart Donaldson at stewart.donaldson@cgu.edu, +1(909)702-7316. The CGU Institutional Review Board (IRB) has certified this project as exempt. If you have any ethical concerns about this project or your rights as a human subject in research, you may contact the CGU IRB at +1(909) 607-9406 or irb@cgu.edu. A copy of this form will be given to you if you wish to keep it.

CONSENT: Selecting “Yes” below means that you understand the information on this form, that someone has answered any and all questions you may have about this study, and you voluntarily agree to participate in it.

- Yes
- No

The undersigned researcher has reviewed the information in this consent form with the participant and answered any of his or her questions about the study.

Signature of Researcher _____________________ Date ___________
Printed Name of Researcher _____________________
Appendix F
Focus Group with Young and Emerging Evaluators Protocol
Qualitative Study of Formal Evaluation Education Programs

Introduction and Consent: Remind that we will be recording the focus group session for transcription purposes and that information about individual or organizational identity will be withheld in reports. The recording will be deleted after the transcription and data cleaning. If any specific details need to be suppressed, please let me know during the focus group session.

Introduction
Young and Emerging Evaluators (YEE) represent the evaluation field’s future workforce. Over the last seven years, the global network EvalYouth has grown and developed, leading to regional and national chapters, with thousands of followers.

The dialogue proposed in this session of Glocal 2022 aims to collect data for research at Claremont Graduate University. The whole study aims to shed light on the discussions about how formal evaluation education programs (FEEP) respond to new challenges in the field of evaluation in a world in transformation. In addition, the discussion aims to explore Young and Emerging Evaluators (YEE) perspectives regarding transformational change and their interests and needs regarding formal evaluation education as a capacity-building strategy.

Section 1
Goals:
1. **Brief presentation** of the participants (15 minutes)
2. **Presentation** of the context and reasoning of the research by Marcia Joppert (10 minutes)

Section 2
**Goal: Plenary Discussion:** What does transformational change mean to Young and Emerging Evaluators? (25 minutes)

Section 3
**Goal: Breakout Group Discussion Instruction – 3 groups of 5, randomly assigned, with use of white board of google docs for summary**
- You have 40 minutes for a group discussion.
- There are three questions (listed below) you can discuss and share your thoughts on.
- After the group discussion, you will be asked to share it back to the plenary.
- Please select one representative (or a facilitator) from the group to share back fundamental discussion or learning points for your group.

1. When looking for opportunities in evaluation education, what are your main expectations?
   a) From your experience, are these expectations easy to fulfill?
b) If not, what are the main barriers YEEs face to accessing the education they need?

2. What should the priorities be for the evaluation education programs to meet your needs?

3. What should evaluation education programs look like to prepare YEE to be involved in transformational change evaluations?

   c) What kind of topics, approaches, and methods are likely to gain more importance in the evaluation practice in the coming years?

   d) Should practicum be required as part of the curriculum?

Group 1 – presents question 1 – the others can complement of comment.
Group 2 – presents question 2 – the others can complement of comment.
Group 3 – presents question 3 – the others can complement of comment.

Section 4

Goal: 30 minutes of discussions

Each group will have 10 minutes to share their group discussion points.

Closing

I am incredibly grateful for the time you have taken with me today. I hope this study will shed light on the discussions about how formal evaluation education programs respond to new challenges in the field of evaluation in a context of a world in transformation, and your contribution today was invaluable.
Appendix G

Draft Email asking for EvalYouth chapters leaders’ support to disseminate the Focus Group
Subject: Focus Group on evaluation training needs and interests
Dear [NAME]
As a Ph.D. student at Claremont Graduate University, I am working on a research project exploring how formal evaluation education programs respond to the challenges in the field of evaluation in a context of a world in transformation. This study will contribute to a comprehensive reflection on formal evaluation education. Participants who volunteer will be asked to participate in a focus group. I would appreciate your support, as an EvalYouth chapter leader, to publicize the focus group information by email or on social networks that have a presence of the YEE in your region/country. The upcoming focus group is planned to be conducted via Zoom meeting, and it is scheduled to take two hours. The event is scheduled to take place on June 2nd, 2022, at 8 am Pacific Standard Time (PST).
To be eligible, participants should:
(a) Be under 35 years of age, or
(b) Be an evaluator who has less than five years of professional experience and has expressed an interest in becoming a professional evaluator within the past five years or less.
Thank you for your time.
Sincerely,
Marcia Joppert, MPA
marcia.joppert@cgu.edu
Ph.D. Student in Evaluation and Applied Methods
Claremont Graduate University
+1(909) 997-0235
Appendix H

Focus Group Consent Form

Agreement to Participate in a Study of Interests and Needs regarding Evaluation Education (IRB # 4164)

You are invited to participate in a research project. By volunteering, you will be asked to participate in a focus group. The focus group aims to gather your perspectives on the role of evaluation to support systems changes and your interests and needs regarding evaluation education. In addition, by participating in this research, you will be helping the evaluation community better understand if current formal evaluation education programs are responding to the challenges in the field of evaluation in a context of a world in transformation.

Volunteering for this study involves no more risk than what a typical person experiences on a regular day. Your involvement is entirely up to you. You may withdraw at any time for any reason. Please continue reading for more information about the study.

Study Leadership: This research project is led by Marcia Joppept of the Claremont Graduate University, who professor Dr. Stewart Donaldson is supervising.

Purpose: This study aims to understand better how formal evaluation education programs respond to the challenges in the field of evaluation in a context of a world in transformation.

Eligibility: To be in this study, you should be under 35 years of age, or an evaluator with less than five years of professional experience and, within the past five years or less, have expressed an interest in becoming a professional evaluator.

Participation: You will be asked to participate in a focus group during the study. The focus group will take about 2 hours of your time.

Risks of Participation: The risks you run by taking part in this study are minimal. The focus group will ask you about your perspectives on the role of evaluation to support systems changes and your interests and needs regarding formal evaluation education programs. Results and participant information will remain confidential. Findings will not be connected with your name.

Benefits of Participation: By participating in the study, you will be contributing to build new evidence regarding YEE’s needs, interests and barriers related to evaluation education. Also, the study will help YEE to identify evaluation education programs responsive to transformational change.

Compensation: You will not be directly compensated for participating in this study.

Voluntary Participation: Your participation in this study is voluntary. You may stop or withdraw from the study or refuse to answer any particular question for any reason at any time without it being held against you. Your decision whether or not to participate will have no effect
on your current or future connection with anyone at CGU. You may contact the CGU IRB at +1(909) 607-9406 or irb@cgu.edu.

Confidentiality: Your individual privacy will be protected in all reports, papers, books, talks, posts, or stories resulting from this study. We may use the data we collect for future research or share it with other researchers, but we will not reveal your identity with it.

Sponsorship: This study is not being sponsored.

Further Information: If you have any questions or want additional information about this study, please contact Marcia Joppert at marcia.joppert@cgu.edu, +1 (909) 997-0235. You may also contact Dr. Stewart Donaldson at stewart.donaldson@cgu.edu, +1(909)702-7316. The CGU Institutional Review Board (IRB) has certified this project as exempt. If you have any ethical concerns about this project or about your rights as a human subject in research, you may contact the CGU IRB at +1(909) 607-9406 or irb@cgu.edu. A copy of this form will be given to you if you wish to keep it.

Consent: Selecting “Yes” below means that you understand the information on this form, you meet the eligibility criteria, that someone has answered any and all questions you may have about this study, and you voluntarily agree to participate in it.

Yes (1)

No (2)
Appendix I
Draft Email Reminder for Focus Group Participants

Subject: Focus Group on evaluation training needs and interests
Dear [YEE name],

You are scheduled for a Focus Group through Zoom meeting on 06.02.2022 at 8 am PST. Please find below the times in each of your locations.

<table>
<thead>
<tr>
<th>Location</th>
<th>Timezone</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claremont, CA, USA*</td>
<td>PDT (UTC -7)</td>
<td>8:00 am</td>
</tr>
<tr>
<td>Mexico City, Mexico*</td>
<td>CDT (UTC -5)</td>
<td>10:00 am</td>
</tr>
<tr>
<td>Toronto, Canada*</td>
<td>EDT (UTC -4)</td>
<td>11:00 am</td>
</tr>
<tr>
<td>São Paulo, Brazil</td>
<td>BRT (UTC -3)</td>
<td>12:00 noon</td>
</tr>
<tr>
<td>Accra, Ghana</td>
<td>GMT (UTC +0)</td>
<td>3:00 pm</td>
</tr>
<tr>
<td>Johannesburg, South Africa</td>
<td>SAST (UTC +2)</td>
<td>5:00 pm</td>
</tr>
<tr>
<td>Maputo, Mozambique</td>
<td>CAT (UTC +2)</td>
<td>5:00 pm</td>
</tr>
<tr>
<td>Harare, Zimbabwe</td>
<td>CAT (UTC +2)</td>
<td>5:00 pm</td>
</tr>
<tr>
<td>Nairobi, Kenya</td>
<td>EAT (UTC +3)</td>
<td>6:00 pm</td>
</tr>
<tr>
<td>Damascus, Syria*</td>
<td>EEST (UTC +3)</td>
<td>6:00 pm</td>
</tr>
<tr>
<td>Delhi, India</td>
<td>IST (UTC +5:30)</td>
<td>8:30 pm</td>
</tr>
<tr>
<td>Thimphu, Bhutan</td>
<td>BTT (UTC +6)</td>
<td>9:00 pm</td>
</tr>
</tbody>
</table>

obs: the times were converted using https://www.timeanddate.com/

Please let me know if you are still available at this time.

The focus group should take approximately 120 minutes and be conducted via zoom through this link: https://us02web.zoom.us/j/86145305590 (please don't share this link with anybody)
The session will be recorded, with your permission, to ensure that my notes are accurate, but recordings will be deleted at the end of this project.

Sincerely,

Marcia Joppert, MPA
marcia.joppert@cgu.edu
Ph.D. Student in Evaluation and Applied Methods
Claremont Graduate University
+1(909) 997-0235
Appendix J

Flyer for Dissemination of the Focus Group in EvalYouth Chapters’ Social Media Channels

WHAT DO YOU THINK EVALUATION EDUCATION SHOULD LOOK LIKE?

WHAT BARRIERS DO YOU FACE?

Please, give us your valuable input on this and other questions by participating in a Focus Group during the GLOCAL-EVAL WEEK 2022
(link to the registration form)
## Appendix K

Code Structure used for Qualitative Analysis - Study 1 and Study 2

<table>
<thead>
<tr>
<th>Parent code</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background</td>
<td>Effects of formal training</td>
</tr>
<tr>
<td></td>
<td>Main contributions to the field</td>
</tr>
<tr>
<td>Transformation and evaluation role</td>
<td>Role of evaluation</td>
</tr>
<tr>
<td></td>
<td>Systemic Thinking</td>
</tr>
<tr>
<td></td>
<td>Transformation</td>
</tr>
<tr>
<td>Implications for the practice</td>
<td>Approaches</td>
</tr>
<tr>
<td></td>
<td>Are these trends becoming evident?</td>
</tr>
<tr>
<td></td>
<td>Changes in power relations</td>
</tr>
<tr>
<td></td>
<td>Competencies</td>
</tr>
<tr>
<td></td>
<td>Contingence on the institutional environment</td>
</tr>
<tr>
<td></td>
<td>Ecosystem building</td>
</tr>
<tr>
<td></td>
<td>Effects of the trends in the interviewee personal practice</td>
</tr>
<tr>
<td></td>
<td>Evaluation spreading into other environments and dis</td>
</tr>
<tr>
<td></td>
<td>Evaluators attitudes and roles</td>
</tr>
<tr>
<td></td>
<td>Methods</td>
</tr>
<tr>
<td></td>
<td>Purpose</td>
</tr>
<tr>
<td></td>
<td>Relevance to public policies</td>
</tr>
<tr>
<td></td>
<td>Stakeholders’ roles</td>
</tr>
<tr>
<td></td>
<td>Training</td>
</tr>
<tr>
<td></td>
<td>Trends (most important themes or topics in the coming years)</td>
</tr>
<tr>
<td>Challenges at a global level in an era of</td>
<td>Advocacy role of evaluators</td>
</tr>
<tr>
<td>transformational changes</td>
<td>Axiological paradigm</td>
</tr>
<tr>
<td></td>
<td>Big data</td>
</tr>
<tr>
<td></td>
<td>Complexity</td>
</tr>
<tr>
<td></td>
<td>Convince people about the need to change</td>
</tr>
<tr>
<td></td>
<td>Culture of evaluation in developing countries</td>
</tr>
<tr>
<td></td>
<td>Culture of Universities and academic system</td>
</tr>
<tr>
<td></td>
<td>Differentiate evaluation from other professional activities</td>
</tr>
<tr>
<td></td>
<td>Diversity of perspectives</td>
</tr>
<tr>
<td></td>
<td>Ecosystem of evaluation - how organizations work together</td>
</tr>
<tr>
<td></td>
<td>Epistemological views</td>
</tr>
<tr>
<td>Parent code</td>
<td>Code</td>
</tr>
<tr>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td>Evaluation learning function</td>
</tr>
<tr>
<td></td>
<td>Evaluators mentality</td>
</tr>
<tr>
<td></td>
<td>Expectations x resources</td>
</tr>
<tr>
<td></td>
<td>Make evaluation useful</td>
</tr>
<tr>
<td></td>
<td>Ontological perspective</td>
</tr>
<tr>
<td></td>
<td>Politics / power relations</td>
</tr>
<tr>
<td></td>
<td>Produce data usable for policy and decision making</td>
</tr>
<tr>
<td></td>
<td>Professionalization of the evaluation</td>
</tr>
<tr>
<td></td>
<td>Project mentality</td>
</tr>
<tr>
<td></td>
<td>Stakeholders’ involvement</td>
</tr>
<tr>
<td></td>
<td>Training</td>
</tr>
<tr>
<td></td>
<td>Transdisciplinarity</td>
</tr>
<tr>
<td></td>
<td>Way evaluations are demanded</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strategies that have been employed given the challenges</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advocacy / Sensitization</td>
<td></td>
</tr>
<tr>
<td>Trust Building</td>
<td></td>
</tr>
<tr>
<td>Changes in data collection methods</td>
<td></td>
</tr>
<tr>
<td>Convince the donors about social justice criteria</td>
<td></td>
</tr>
<tr>
<td>Create a learning environment</td>
<td></td>
</tr>
<tr>
<td>Discussions and publications on evaluators profile/competence</td>
<td></td>
</tr>
<tr>
<td>Empower actors</td>
<td></td>
</tr>
<tr>
<td>Negotiate feasible designs</td>
<td></td>
</tr>
<tr>
<td>People or organizations trying to build the bridge</td>
<td></td>
</tr>
<tr>
<td>Produce knowledge resources</td>
<td></td>
</tr>
<tr>
<td>Proposal/Use of different approaches</td>
<td></td>
</tr>
<tr>
<td>Role for international organizations</td>
<td></td>
</tr>
<tr>
<td>Use of international or new frameworks</td>
<td></td>
</tr>
<tr>
<td>VOPEs role to influence demand and donors</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Barriers or obstacles</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgement of the relevance of evaluations</td>
<td></td>
</tr>
<tr>
<td>Difficulty to implement new approaches</td>
<td></td>
</tr>
<tr>
<td>Evaluation capacity</td>
<td></td>
</tr>
<tr>
<td>Lack of data</td>
<td></td>
</tr>
<tr>
<td>Lack of resources</td>
<td></td>
</tr>
<tr>
<td>Mental models and existing practice</td>
<td></td>
</tr>
<tr>
<td>Organizational resistance</td>
<td></td>
</tr>
<tr>
<td>Political will</td>
<td></td>
</tr>
<tr>
<td>Time pressure</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Implications for evaluation education</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application of different theories in a transformative environment</td>
<td></td>
</tr>
<tr>
<td>Parent code</td>
<td>Code</td>
</tr>
<tr>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>Approaches / theories</td>
<td></td>
</tr>
<tr>
<td>Beliefs about the importance of formal training</td>
<td></td>
</tr>
<tr>
<td>Connection with competencies and skills needed</td>
<td></td>
</tr>
<tr>
<td>Customize training initiatives with context needs</td>
<td></td>
</tr>
<tr>
<td>Education Ecosystem / partnerships</td>
<td></td>
</tr>
<tr>
<td>Evaluation as a discipline or a transdiscipline</td>
<td></td>
</tr>
<tr>
<td>Evaluation as a professional practice (+)</td>
<td></td>
</tr>
<tr>
<td>Learning Assessment associated to Knowledge Creation</td>
<td></td>
</tr>
<tr>
<td>Level / type of degree</td>
<td></td>
</tr>
<tr>
<td>Limitations of evaluation education</td>
<td></td>
</tr>
<tr>
<td>Mentorship</td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td></td>
</tr>
<tr>
<td>Modality</td>
<td></td>
</tr>
<tr>
<td>Practicum experience / internship</td>
<td></td>
</tr>
<tr>
<td>Principles</td>
<td></td>
</tr>
<tr>
<td>Skills</td>
<td></td>
</tr>
<tr>
<td>Target groups</td>
<td></td>
</tr>
<tr>
<td>Teaching Methods</td>
<td></td>
</tr>
<tr>
<td>Topics</td>
<td></td>
</tr>
</tbody>
</table>