Journal of Interreligious Studies
July 2025, Issue 46, 88–109
ISSN 2380-8187
www.irstudies.org

ARTICLE

The (Im)Possibility of Measuring Interfaith Learning Outcomes

Hannah J. Visser, Gerdien D. Bertram-Troost, Marianne Moyaert, and Anke I. Liefbroer

Abstract

Despite the growing number of quantitative studies on interfaith learning outcomes, any reflection on the motivations and implications of such measurement has remained notably absent. This article argues that such reflection is necessary if we do not want quantitative measurement to misrepresent or even impair what is happening in interfaith initiatives. Drawing on debates in neighboring disciplines such as the educational sciences and peacebuilding studies, the article explores why we should—or should not—measure these outcomes. It then identifies other possible ways of measuring interfaith learning outcomes that still do justice to the complex reality of interfaith learning. Based on interdisciplinary literature, it suggests that at least three imperatives should be taken into consideration when measuring interfaith learning outcomes: 1) involve multiple stakeholders, 2) describe and include the context, and 3) use methods other than, and in addition to, quantitative measurement. The implications and challenges of these suggestions are explored through our own experience of evaluating the outcomes of an interfaith leadership program. The article demonstrates how quantitative measurement can provide helpful insights but cannot capture all aspects of the messy, complex, and multifaceted nature of interfaith learning. It is intended to stimulate reflection on the why and how of measuring "what works."

Keywords

interfaith learning, quantitative measurement, evaluation, learning outcomes

In light of the rapid increase in interfaith initiatives over the past few decades, interfaith scholars and practitioners have questioned whether these initiatives are achieving their goals. They have asked themselves which activities are most effective in bringing about the desired change. To answer these questions, a growing number of studies have recently considered which quantitative instruments are best suited to assess the impact of interfaith learning;[1] there have also been several recent attempts to quantitatively evaluate interfaith initiatives.[2] However, although most of these studies have focused on the *outcomes* of interfaith learning and the *empirical tools* used to assess these outcomes, there has been a remarkable lack of reflection on the *underlying assumptions* and *practical implications* of quantitatively measuring these learning outcomes. We argue here that such reflection is necessary if we do not want quantitative measurements to misrepresent or even impair what is happening in interfaith initiatives.

The assumptions and challenges of measuring learning outcomes and its impact on educational practice have been fiercely debated in neighboring disciplines such as educational sciences and peacebuilding studies.[3] While

1 See, for example, Alyssa N. Bryant, "Assessing Contexts and Practices for Engaging Students' Spirituality," *Journal of College and Character* 10, no. 2 (2008); The Woolf Institute, *How to Measure Success: A toolkit for the evaluation of interfaith engagement* (2021); Matthew J. Mayhew and Alyssa N. Rockenbach, "Interfaith Learning and Development," *Journal of College and Character* 22, no. 1 (2021/02/01 2021). Also see the IDEALS survey (https://www.interfaithamerica.org/research/ideals/) and the INSPIRES Index (https://www.inspiresindex.org/)

2 See, for example, Assumpta Aneas and Ruth Vilà, "Evaluation of Intercultural and Interreligious Dialogue Competencies. Identification of Factors Related with Its Performance among Adolescents in the City of Barcelona," *Religion & Education* (2023); Evelyn Schnaufer et al., "Developing a Research Tool for Investigating Religious Knowledge as Part of Religious Literacy: The Questionnaire – First Results – Possibilities for International Comparisons," *Journal of Empirical Theology* 35, no. 2 (2023); Lucy Peacock, "Contact-based interfaith programmes in schools and the changing religious education landscape: negotiating a worldviews curriculum," *Journal of Beliefs & Values* 44, no. 1 (2023).

3 See, for example, Izaak Dekker and Martijn Meeter, "Evidence-based education: Objections and future directions," Conceptual Analysis, *Frontiers in Education* 7 (2022); Mohammed Abu-Nimer and Renáta Katalin Nelson, eds., *Evaluating Interreligious Peacebuilding and Dialogue: Methods and Frameworks*, KAICIID – Beyond Dialogue Series (Berlin, Boston: De Gruyter, 2021); Gerdien D. Bertram-Troost, "Investigating the impact of religious diversity in schools for secondary education: a challenging but necessary exercise," *British Journal of Religious Education* 33, no. 2 (2011); Tobias Denskus, "Challenging the international peacebuilding evaluation discourse with qualitative methodologies," *Evaluation and Program Planning* 35, no. 1 (2012).

measurement is commonly accepted as standard practice in these fields, it has also faced increasing criticism, such as Biesta's critique of the "age of measurement" in education, or Van der Kolk's reference to the individual and societal "costs" of performance measurement.[4] It is our impression that these debates have not yet been engaged in the emerging field of interfaith learning even though the call for measurement is sounding louder here too. With the rapid increase of quantitative measurements of interfaith learning outcomes, the field now appears to be at a crossroads, and we must choose between prioritizing quantitative outcomes and evidence-based practice on the one hand, or moving towards more narrative, qualitative approaches to understanding the processes of interfaith learning on the other. We therefore believe that now is the time to start reflecting on the assumptions and challenges of measuring interfaith learning outcomes.

The purpose of this article is not to advocate for or against measuring interfaith learning outcomes. Rather, it is intended to initiate a conversation among interfaith program organizers and researchers in the field of interfaith learning about the underlying assumptions and practical challenges of measuring interfaith learning outcomes. The terms "measure" and "measurement" in this article refer to measuring the outcomes of an initiative using *quantitative* methods. When discussing "evaluation," we refer to the broader process of trying to understand participants' experiences, learning outcomes, and the impact of an interfaith initiative.[5] An "interfaith initiative" is understood as an organized activity that intentionally and actively brings together individuals or communities who orient around religion differently for a variety of learning objectives.[6] These activities can be organized in (post-)conflict societies, thus overlapping with the field of (inter-)religious peacebuilding studies.[7] However, they can also be organized

4 Regarding Biesta's critique, see *Good Education in an Age of Measurement*, ed. Michael A. Peters and Colin Lankshear, Interventions: Education, Philosophy & Culture, (New York: Routledge, 2010). Regarding the costs of performance measurement, see "Numbers Speak for Themselves, or Do They? On Performance Measurement and Its Implications," *Business & Society* 61, no. 4 (2022).

5 See Donald L. Kirkpatrick and James D. Kirkpatrick, *Evaluating Training Programs: The Four Levels*, Third edition ed. (San Francisco: Berrett-Koehler Publishers, Inc., 2006).

6 Regarding this working definition of an interfaith initiative, see Kusumita P. Pedersen, "The interfaith movement: an incomplete assessment," *Journal of Ecumenical Studies* 41, no. 1 (Winter 2004): 79. Regarding learning objectives, see Hannah J. Visser and others, "Categorising Interfaith Learning Objectives: A Scoping Review," *Journal of Beliefs & Values* 44, no. 1 (2023).

7 For example, Mohammed Abu-Nimer, "Conflict Resolution, Culture, and Religion:

in educational settings in other contexts, such as the influential initiatives on U.S. campuses. In this article, the term "field of interfaith learning" refers to the body of scholarship surrounding interfaith initiatives. Our work draws on three key subfields: 1) *interreligious/interfaith studies*, which has made significant contributions to the organization and evaluation of interfaith initiatives, though building a strong empirical foundation remains a challenge;[8] 2) *religious education*, which has focused more extensively on empirical research, particularly on fostering students' beliefs, religious literacy, and attitudes towards religious diversity in institutional education settings;[9] and 3) *(inter) religious peacebuilding*, which examines how religious actors can support or hinder sustainable peace efforts, along with methods for evaluating these initiatives.[10]

In the next section, we discuss the question of *why* (not) to use quantitative measures to evaluate interfaith initiatives. We will draw on debates from related disciplines to provide insight into this question. Based on this interdisciplinary literature, we propose taking into account several suggestions when measuring interfaith learning outcomes while also addressing critiques of such measurement. In the second part of this article, we will discuss *how* to measure interfaith learning outcomes in a way that reflects the complex realities of interfaith learning. As we, too, have attempted to evaluate an interfaith initiative, we will draw on our own experiences of designing and applying an evaluation approach, including a pilot measurement tool, to illuminate these underlying challenges and tensions.

Toward a Training Model of Interreligious Peacebuilding," *Journal of Peace Research* 38, no. 6 (2001).

8 For example, Rachel Mikva, *Interreligious Studies: An Introduction* (Cambridge: Cambridge University Press, 2023); Hans Gustafson, ed., *Interreligious studies: dispatches from an emerging field* (Waco: Baylor University Press, 2020); Eboo Patel, Jennifer Howe Peace, and Noah J. Silverman, eds., *Interreligious/ Interfaith Studies: Defining a New Field* (Boston: Beacon Press, 2018).

9 For example, Friedrich Schweitzer et al., "Current Debates about (Inter-) Religious Literacy and Assessments of the Outcomes of Religious Education: Two Approaches to Religion-Related Knowledge in Critical Review," *Journal of Beliefs & Values* 44, no. 2 (2023).

10 For example, Peter Woodrow, Nick Oatley, and Michelle Garred, *Faith Matters: A Guide for the Design, Monitoring & Evaluation of Inter-Religious Peacebuilding*, CDA Collaborative Learning Projects and Alliance for Peacebuilding (2017).

1. Why (Not) Measure Interfaith Learning Outcomes?

Various scholars and practitioners in the field of interfaith learning have emphasized the importance of evaluating interfaith initiatives, particularly through quantitative measurement tools.[11] According to them, measuring the outcomes of interfaith learning can tell us more about "what works." This knowledge could help to improve practices and attract more funding when potential funders gain more insight into the foreseen outcomes of an interfaith initiative. Relying solely on narrative or qualitative approaches does not seem to be sufficient, as these approaches are seen as subjective and incomparable to other contexts, whereas "a more robust design…can increase the ability of evaluators to accurately assess results and enhance the credibility of the final product."[12] Before discussing the arguments against measuring interfaith learning outcomes, we will first examine the arguments in favor of quantitative measurement.

1.1. Motivations for Measuring Interfaith Learning Outcomes

Scholars in various disciplines, but particularly in the field of peacebuilding, generally cite two main motivations for measuring program outcomes: learning and accountability.[13] The "learning" argument suggests that measuring interfaith learning outcomes can provide insight into what the program is achieving, which can guide improvement of the program. This argument appeals to the intrinsic motivation of organizers and facilitators of interfaith initiatives, who seek concrete feedback to improve the program's relevance and impact. Advocates of using quantitative measurement to

11 For example, Renee Garfinkel, *What works? Evaluating interfaith dialogue programs*, United States Institute of Peace (Washington DC, 2004); Jennie Vader, *Meta-review of inter-religious peacebuilding program evaluations*, CDA Collaborative Learning Projects (2015); Conor Seyle et al., *Some Credible Evidence: Perception about the Evidence Base in the Peacebuilding Field*, One Earth Future Foundation; The Alliance for Peacebuilding (2021); Richard McCallum, "Towards a framework and methodology for the evaluation of inter-faith initiatives," *Studies in Interreligious Dialogue* 27, no. 1 (2018).

12 Vader, *Meta-review of inter-religious peacebuilding program evaluations*, 38.

13 For example, Mohammed Abu-Nimer and Renáta Katalin Nelson, "Introduction: Evaluating Religious and Interreligious Peacebuilding: Meeting the Challenge," in *Evaluating Interreligious Peacebuilding and Dialogue. Methods and Frameworks*, ed. Abu-Nimer Mohammed and Nelson Renáta Katalin (Berlin, Boston: De Gruyter, 2021); Seyle et al., *Perception about the Evidence Base in the Peacebuilding Field*.

evaluate interfaith learning argue that this information ultimately benefits participants, who thereby take part in a relevant and effective program.[14]

More specifically, quantitative measurement has the potential to facilitate learning by providing standardized measurements that can be compared across groups. This can be done in several ways, such as comparing pre- and post-program outcomes over time within the same context in order to assess the impact of an initiative on participants' competencies, comparing initiatives in a similar context to understand progress and differences, or comparing initiatives and practices in different contexts to help identify best practices.[15] Using such standardized approaches, the idea is that "we could see genuine, generational progress instead of the usual pendulum swings of opinion and fashion."[16]

In addition to the motivation to learn and improve, there is another reason that is often given: the 'accountability' argument.[17] This argument claims that the quantitative measurement of program outcomes is necessary in order to prove an initiative's worth, both literally and figuratively.[18] Interfaith initiatives, which are often dependent on donor funding or accreditation, are now required to provide tangible results on whether and to what extent they are bringing about the intended changes. Donors and policymakers also seek positive impact with limited resources, and "the reality of conducting interfaith engagement in the modern world is that it almost always relies on funding. And funders rely, to an extent, on evidence of impact, and answers to questions such as: Has the money been well-spent? Was the grant given to the right organization?"[19] One could see this motivation for measuring as a loss of autonomy and as a willingness to conform to the normative ideas of funders and policymakers.[20] However,

14 Friedrich Schweitzer, "Researching classroom processes and outcomes in religious education: the need for intervention studies," in *Researching religious education: classroom processes and outcomes*, ed. Friedrich Schweitzer and Reinhold Boschki (Münster: Waxmann, 2018), 196.

15 Consider Philip Davies, "What is evidence-based education?," *British journal of educational studies* 47, no. 2 (1999): 114.

16 Robert E. Slavin, "Evidence-based education policies: Transforming educational practice and research," *Educational researcher* 31, no. 7 (2002): 20.

17 Abu-Nimer and Nelson, "Introduction."

18 Hippolyt Pul, "When My Peace Is Not Your Peace," in *Evaluating Interreligious Peacebuilding and Dialogue: Methods and Frameworks*, ed. Mohammed Abu-Nimer and Renáta Katalin Nelson (Berlin, Boston: De Gruyter, 2021).

19 The Woolf Institute *A toolkit for the evaluation of interfaith engagement*, 38.

20 See Gerdien D. Bertram-Troost, Menswording in een laag-vertrouwensamenleving: Kansen en uitdagingen voor onderwijs., 2022, Vrije Universiteit Amsterdam; Reina

despite these critiques of the "era of accountability",[21] as we will discuss below, the perceived need for quantitative evaluation tools to measure interfaith learning outcomes is understandable because it can attract the necessary financial support. Thus, from this perspective, a first step in professionalizing the field and attracting funding would be to develop and apply more quantitative tools.[22] Quantitative measurement can help to prioritize this work for funders and policymakers and to underscore its importance in building peaceful societies.

Both these arguments seem understandable and at first sight more or less convincing: who would not want to improve initiatives and show the results of interfaith learning? However, on closer inspection of ongoing discussions about measurement in neighboring disciplines, it becomes clear that these arguments are not self-evident.

1.2. Reasons to Be Cautious About Measuring Interfaith Learning Outcomes

To explore the reasons for caution in measuring interfaith learning outcomes, we will focus on three interrelated concerns. The first concern, which has been raised primarily in debates in the educational sciences, is that measurement can contribute to a pervasive 'culture' of prioritizing accountability and results.[23] The second and third concerns are often raised in discussions of measurement and evaluation in the field of (religious) peacebuilding as well: the challenge or even impossibility of measuring intangible learning outcomes and the difficulty of doing justice to a specific context.[24]

C. Neufeldt, "Vying for Good," in *Evaluating Interreligious Peacebuilding and Dialogue: Methods and Frameworks*, ed. Mohammed Abu-Nimer and Renáta Katalin Nelson (Berlin, Boston: De Gruyter, 2021).

21 Jill Anne Chouinard, "The Case for Participatory Evaluation in an Era of Accountability," *American Journal of Evaluation* 34, no. 2 (2013).

22 For example, Bryant, "Assessing Contexts and Practices for Engaging Students' Spirituality."; Jeremy T. Snipes and Benjamin Correia-Harker, "Revisiting the Assessment Context: A Call to Interfaith Assessment," *Journal of College and Character* 18, no. 2 (2017/04/03 2017); Vader, *Meta-review of inter-religious peacebuilding program evaluations*; The Woolf Institute, *A toolkit for the evaluation of interfaith engagement*, 38.

23 For example, Biesta, *Good Education in an Age of Measurement*; Van Der Kolk, "On Performance Measurement and Its Implications;" Chouinard, "The Case for Participatory Evaluation in an Era of Accountability."; Bertram-Troost, Menswording in een laag-vertrouwensamenleving.

24 For example, Woodrow, Oatley, and Garred, *Faith Matters*; Pul, "When My Peace Is Not Your Peace."

The first and most substantial concern about measuring learning outcomes is directed towards the 'culture' of accountability and evidence. As educational philosopher Gert Biesta has warned, measurable learning outcomes are often assumed to be the only and most important indicator of learning.[25] According to him, measurement narrows the understanding and organization of educational practice. It leads to a situation where measurable outcomes are valued more than the nuanced, intangible educational achievements. This focus on measurable results can make measurement an end in itself, whereby "we are just measuring what we can easily measure and thus end up valuing what we (can) measure."[26] This situation can have significant implications for educational practice, which may then be organized to prioritize the attainment of measurable outcomes over other considerations of what entails "good" education.[27] The concept of "indicatorism" sheds light on the potential negative consequences of this situation:

> Overly stressing the importance of performance measures may trigger behavior that is aimed at improving performance indicators, while losing sight of the actual goals. I call this *indicatorism*. (…) Indicatorism also includes improving an indicator for the short term (at the expense of the long term), ignoring "unmeasured" tasks, and manipulating numbers.[28]

Consequently, the centralization of quantitative measurement can have a detrimental impact on the practice of interfaith initiatives, diverting attention, time, and funding away from what the values and ideals of teachers and organizers consider to be 'good,' and toward the outcomes of interfaith initiatives that are merely measurable.

A second reason for caution in measuring interfaith learning outcomes is the difficulty, if not impossibility, of accurately representing the reality of interfaith work in a standardized way. The process of measuring the outcomes of such a program is "intrinsically complex" because it is

25 Biesta, *Good Education in an Age of Measurement*.

26 Biesta, *Good Education in an Age of Measurement*, 13; For a reflection on these dynamics in the context of (inter)religious education, also see Bertram-Troost, *Menswording in een laag-vertrouwensamenleving*.

27 Biesta, *Good Education in an Age of Measurement*.

28 Van Der Kolk, "On Performance Measurement and Its Implications," 814–15. Emphasis in original.

inherently intertwined with the lived experiences of participants.[29] This complexity is further compounded by the fact that interfaith initiatives focus on outcomes that are difficult to measure, such as mutual appreciation, leadership, and self-awareness.[30] Due to the intangible nature of interfaith learning outcomes, and the additional observation that these outcomes are defined and understood differently by each individual, the risk that a quantitative measurement will not accurately measure what it is intended to measure is considerable.[31] As Neufeldt describes in her analysis of different positions in the debate around measurable peacebuilding indicators, critics often argue that because peacebuilding programs are inherently intangible, messy, and highly contextual, any attempt to standardize outcomes is doomed to failure. Instead, the critics are "interested in the uniqueness of interventions and communities—they focus on the stories and lessons that emerge from specific cultural, geographic and temporal contexts and do not expect these to be generalizable."[32]

A third concern regarding the measurement of learning outcomes is that it may fail to account for the changing and multifaceted contexts in which interfaith learning occurs. The socio-political factors that influence initiatives make it difficult to capture anything other than the outcomes at a specific point in time and in a particular place.[33] Each initiative is situated within a distinct context and timeframe, characterized by specific attributes. These include recent political developments (for example, the rise of far-right political parties in our Western European context), historical legacies of colonialism, religious hegemony and oppression, and national legal

29 Jennifer C. Greene, "The Inequality of Performance Measurements," *Evaluation* 5, no. 2 (1999): 163.no. 2 (1999

30 Hannah J. Visser, "Evaluating the Impact of Interfaith Learning: Definitions and Challenges " in *From Interreligious Learning to Interworldview Education*, ed. Didier Pollefeyt, Bibliotheca Ephemeridum Theologicarum Lovaniensium (Leuven: Peeters Publishers, 2023).

31 Khaled Ehsan, "Exploring Power Dynamics of Religious Leaders: The Need for Objectivity," in *Evaluating Interreligious Peacebuilding and Dialogue. Methods and Frameworks*, ed. Mohammed Abu-Nimer and Renáta Katalin Nelson, KAICIID—Beyond Dialogue Series (Berlin, Boston: De Gruyter, 2021), 129; Bertram-Troost, "Investigating the impact of religious diversity in schools," 274.

32 Reina C. Neufeldt, "'Frameworkers' and 'Circlers'—Exploring Assumptions in Impact Assessment," *Advancing conflict transformation: The Berghof handbook II* (2011): 486.

33 Mohammed Abu-Nimer, "Challenges in Peacebuilding Evaluation. Voices from the Field," in *Evaluating Interreligious Peacebuilding and Dialogue. Methods and Frameworks*, ed. Mohammed Abu-Nimer and Renáta Katalin Nelson, KAICIID—Beyond Dialogue Series (Berlin, Boston: De Gruyter, 2021).

frameworks of religion and secularism. So, whereas comparison is one of the key objectives of measurement, it is inherently difficult to compare the outcomes of interfaith initiatives in different contexts or at different points in time because these outcomes are influenced by such a wide range of political, historical, and legal factors. Consequently, the outcomes of interfaith initiatives cannot be attributed solely to the pedagogical approach of an initiative itself.[34] This concern raises the question of whether quantitative measurement provides any useful information at all.

1.3. Suggestions When Measuring Interfaith Learning Outcomes

In sum, certain aspects of interfaith learning are intangible and difficult to measure.[35] This may seem self-evident, but as critiques by scholars like Biesta show, there is a pervasive tendency to present and interpret quantitative measurement as the essential truth of what is happening in education.[36] These critiques raise awareness of the potentially harmful tendencies of quantitative measurement and emphasize the importance of recognizing the unmeasurable aspects and the contexts of the people involved. However, while it is crucial to heed these critiques, it is equally important not to dismiss the merits of quantitative measurement entirely; some of the concerns about measuring interfaith learning outcomes can be avoided, at least to some extent, by taking certain considerations into account. Quantitative measurement, when implemented correctly, can provide valuable information about participants' development over time, help compare different approaches and activities, and enhance the visibility of interfaith work among policymakers and funders. The solution, then, may be not to abandon measurement altogether or to rely exclusively on it, but rather to "insist that definitions and measures of program quality honor the inherent complexity, plurality, and dynamic intersubjectivity of human experience."[37]

Drawing on the concerns and recommendations discussed above, we propose that *if* one wants to quantitatively measure interfaith learning

34 McCallum, "Towards a framework and methodology for the evaluation of inter-faith initiatives," 71; Friedrich Schweitzer, "Religionsunterricht erforschen: Aufgaben und Möglichkeiten empirisch-religionsdidaktischer Forschung," *Zeitschrift für Pädagogik und Theologie* 60, no. 1 (2008): 72; Bertram-Troost, "Investigating the impact of religious diversity in schools."

35 Visser, "Evaluating the Impact of Interfaith Learning."

36 Biesta, *Good Education in an Age of Measurement*.

37 Greene, "The Inequality of Performance Measurements," 171.

outcomes, one should consider these three suggestions: 1) involve multiple stakeholders, 2) include the context, and 3) use other methods in addition to quantitative measurement. We will now briefly explain each of these suggestions before discussing their practical implementation in the second part of this article.

The first suggestion, involving multiple stakeholders, addresses the risk of having quantitative measurements "assess" participants' performance or facilitators' responsibility without considering their experiences and perspectives. Quantitative measurements often focus on indicators defined by researchers or evaluators who have administered a final questionnaire to participants.[38] These indicators may not always reflect the lived reality of the participants or other stakeholders. This gap can lead to (a double) misunderstanding, with participants interpreting researchers' questions differently than intended, and vice versa.[39] As Pul notes, this issue concerns the validity of a measurement instrument—that is, whether it accurately measures what it intends to measure—but it also involves the ownership of defining what is to be measured.[40] Involving different groups of participants and facilitators in identifying what constitutes "success" ensures that measurement reflects each of their realities as much as possible.[41]

Secondly, quantifying learning and its outcomes inevitably reduces a complex reality to measurable metrics, often neglecting the specific context in which a program is organized, whereas "politics and community dynamics play crucial roles in every aspect of the program."[42] As a response, many researchers and evaluators emphasize using an approach that is more sensitive to the way contextual factors shape the experiences of those involved.[43] Context-sensitive measurement examines a program's context in addition to its outcomes, which is especially important in interfaith learning initiatives addressing societal challenges. Understanding the context is essential to comprehending learning outcomes such as leadership skills, contextual knowledge, and awareness of injustice.[44]

38 Pul, "When My Peace Is Not Your Peace."

39 Bertram-Troost, "Investigating the impact of religious diversity in schools," 277; Saville Kushner, *Personalizing Evaluation* (London2000), 10.

40 Pul, "When My Peace Is Not Your Peace," 78. cf. Neufeldt, "Vying for Good," 56.

41 Bertram-Troost, "Investigating the impact of religious diversity in schools."

42 Garfinkel, *Evaluating interfaith dialogue programs*, 22.

43 Jeremy T. Snipes and Benjamin P. Correia-Harker, "Implementing Inclusive Interfaith Assessment: Considerations and Challenges," *New Directions for Student Services* 2020, no. 169 (03/01/2020): 22.

44 Visser et al., "Categorising Interfaith Learning Objectives: A Scoping Review."

The third suggestion we would like to consider is to use multiple methods—that is, not to rely solely on quantitative methods. Scholars in educational sciences, peacebuilding literature, and interfaith studies tend to advocate for the inclusion of qualitative methods such as interviews or focus groups in addition to quantitative methods.[45] There are several reasons for bringing together qualitative and quantitative methods, such as data-triangulation, complementing insights, and using outcomes of one method to develop another.[46] Overall, this approach captures both a breadth of data on the differences between programs or groups and participants' learning over time, as well as a depth of insight into the "when, why and how" behind changes (or the lack thereof) in participants' attitudes and behaviors.[47] Thus, using multiple methods to understand the process and outcomes of an interfaith program allows for a more comprehensive view than relying on just one method.

2. How (Not) to Measure Interfaith Learning Outcomes

In this second part of our article, we discuss how, in practice, to measure interfaith learning outcomes based on the three suggestions proposed above. Our insights stem from our experience with a study evaluating an interfaith leadership program in Belgium and the Netherlands. We will begin with an introduction to the research project.

In 2020, we started a research project on the impact of interfaith learning, focusing on the Emoena interfaith leadership program in Belgium and the Netherlands. Originally launched in France in 2015, Emoena has been organized in Belgium and the Netherlands since 2019. This year-long program aims to enhance the interfaith leadership of professionals by improving dialogue skills, expanding (inter)religious literacy, developing leadership skills, increasing knowledge about societal challenges, and training participants to organize community projects.[48] Our study analyzed two

45 Abu-Nimer, "Challenges in Peacebuilding Evaluation," 33.

46 Jennifer C. Greene, Valerie J. Caracelli, and Wendy F. Graham, "Toward a Conceptual Framework for Mixed-Method Evaluation Designs," *Educational Evaluation and Policy Analysis* 11, no. 3 (1989).

47 David Steele and Ricardo Wilson-Grau, "Transcendence and the Evaluation of Faith-Based Peacebuilding," in *Evaluating Interreligious Peacebuilding and Dialogue: Methods and Frameworks*, ed. Mohammed Abu-Nimer and Renáta Katalin Nelson (Berlin/Boston: De Gruyter, 2021), 153.

48 See: Hannah J. Visser et al., "Learning Orientations in Interfaith Initiatives: A Case Study of the Interfaith Leadership Program Emoena," *Religious Education* 118, no. 4 (2023): 371–72.

cohorts in Belgium and the Netherlands during the 2021–2022 academic year using a mixed methods approach and included three methods.[49] First, we analyzed participants' reflection reports during their participation in Emoena in the Netherlands in order to deepen our understanding of the interfaith learning process.[50] Second, we interviewed participants at the end of the Emoena program to understand the way in which they evaluated their experiences, learning outcomes, transfer, and ideas for the future.[51] Third, we conducted a pilot quantitative study with seventeen participants, to measure their skills and societal awareness.

For the quantitative pilot study, we collaborated with participants, facilitators, and interfaith experts to develop case stories that presented real-life dilemmas related to interfaith leadership. Unlike common attitude-scales (for example, "How open are you towards Buddhists, on a scale of one to ten?"), these scenarios required action-based responses, reflecting the complexity of real-life situations.[52] Participants responded to eight case stories with closed-ended options linked to indicators such as "listening to others" or "self-reflection on biases."[53] We measured changes in participants' skills and awareness by collecting responses before, immediately after, and six months after their participation in the Emoena program. An example case story goes as follows:

> During a day on worldview for high school students, people with different worldviews were invited to come and talk about their perspectives. They took turns talking about their worldview. Leon, a presenter from a Reformed Christian background, presents his worldview first. Next, Sita speaks about rituals and customs within Indian Hinduism. Leon begins to sigh deeply. He gets up and leaves the classroom

49 See www.emoena.be for a description of the Emoena Belgium program and www.vu.nl/emoena for a description of the Emoena Netherlands program.

50 The methods and results of this approach have been published elsewhere: Visser et al., "Learning Orientations in Interfaith Initiatives."

51 Kirkpatrick and Kirkpatrick, *Evaluating Training Programs*. The methods and results of the interview study have been conditionally accepted for publication elsewhere: Hannah J. Visser et al., "Beyond Interfaith Harmony: How Participants Perceive the Impact of an Interfaith Leadership Program in Flanders and the Netherlands," (forthcoming).

52 This approach was inspired by the vignette study conducted by Martin Losert, Heinrich Merkt, and Friedrich Schweitzer, "In Search of Interreligious Competence: An Empirical Study in the Context of Training Caregivers Through Religious Education," *Journal of Empirical Theology* 28, no. 1 (2015).

53 Visser et al., "Categorising Interfaith Learning Objectives: A Scoping Review."

without saying anything. Only after Sita's presentation is over, during a short break, does he re-enter the classroom. As the teacher, you are responsible for this day. You stay in your seat and see Leon again during the break. In this situation, would you do the following? (Mark each possibility: 1 = definitely not; 2 = probably not; 3 = not sure/neutral; 4 = probably; 5 = definitely)

1. I would talk to the students about how these encounters are sometimes confronting and discuss their importance;

2. I would consider whether any inequalities between Sita and Leon come into play in this situation;

3. I would consider how I feel about this situation;

4. I would take Leon aside and ask if he would tell me why he left the classroom;

5. I would ask myself if there was anything I could have done differently prior to the day to prevent this from happening;

6. I would ask Leon to leave altogether because he is not contributing to the purpose of this day (reverse scaled item).

We will now turn to further reflection on the practical implications and challenges of measuring interfaith learning outcomes according to the three considerations described above, using our research project as an illustration of the practical challenges and opportunities.

2.1. Involving Multiple Stakeholders

In our pilot study, we involved multiple stakeholders for several reasons: it not only enhances the validity of quantitative measurements but also fosters a broader sense of ownership regarding the definition of "success."[54] Given the diverse definitions and interpretations of interfaith learning outcomes among stakeholders, there is always a risk of misinterpretation. Stakeholder

54 Pul, "When My Peace Is Not Your Peace."

involvement underscores the importance of bridging the gap among researchers, facilitators, and participants.

In practical terms, stakeholder involvement can range from full collaboration throughout the process to consultation at specific stages. Cousins and Whitmore identify three dimensions of such collaboration. First, control over the evaluation process can rest solely with the researchers or be shared with a diverse group of practitioners. Second, the scope of stakeholder involvement can vary, including only the primary "users"—such as program participants—or expanding to encompass a broader range of potential stakeholders. Third, the level of participation can differ, from minimal consultation to deep, continuous engagement throughout the evaluation process.[55] Through our research project, we gained insight into this principle. We began by analyzing participants' reflection reports from the Emoena program and conducting interviews with program coordinators, gaining insights into their perspectives on successes and challenges. For instance, our analysis revealed diverse expectations and orientations for interfaith learning, ranging from gaining theoretical knowledge to developing concrete competencies for organizing initiatives.[56] Consequently, we concluded that interfaith learning outcomes would be bound to vary among participants and could not be held to a single standard. Such input could be used for interpreting our quantitative outcomes, as it highlights potential differences between participants.

We integrated these perspectives from participants and coordinators into the development of our measurement tool, which includes brief case stories of lived interfaith leadership. To ensure broader stakeholder involvement, we invited experts in the Dutch interfaith landscape to respond to the case stories, guiding the formulation of measurement items. This process allowed us to identify indicators of successful or challenging interfaith encounters that we might have overlooked using solely an academic or external lens. For instance, it highlighted different interpretations of "leadership"; some stakeholders focused on organizational skills, while others emphasized personal leadership.

Moreover, engaging multiple stakeholders can enhance the validity of a measurement tool, ensuring it measures what it intends to measure. Scholars recommend involving participants in a pre-test or pilot-test. This is especially important in fields where terms such as "religion" or "worldview" are

55 J. Bradley Cousins and Elizabeth Whitmore, "Framing participatory evaluation," *New Directions for Evaluation* 1998, no. 80 (1998).

56 Visser et al., "Learning Orientations in Interfaith Initiatives."

subject to diverse interpretations.[57] As Bertram-Troost notes in the context of research on religious diversity in secondary schools: "[I]n the whole research process, methods also need to be found to check how questions are interpreted by pupils and to ascertain what pupils think about the ways researchers interpret their answers. Therefore, it is very important to make use of a pilot test prior to the final data collection. A short oral interview or a group discussion with the persons who filled in the questionnaire should be part of this."[58] This recommendation applies to other populations as well. In the development of our measurement tool, we included a pre-test with participants through post-interview interviews, where participants provided feedback on their understanding of the items immediately after completing the measurement tool. This feedback was incorporated to make the tool understandable and realistic for participants.[59]

By involving multiple stakeholders throughout the measurement process, from conceptualization to validation, we learned about the perspectives on, and experiences with, interfaith leadership of the participants with whom we interacted, ensuring that they understood our questions as we intended them. This collaborative approach enhances inclusivity and validity when measuring program outcomes and defining indicators of "success." It enhances the credibility of research findings and promotes a deeper understanding of the realities experienced by participants and facilitators in interfaith contexts.

2.2. Context Sensitivity

When measuring the outcomes of interfaith learning, context sensitivity is paramount, as it reflects changes in participants behaviors within their own sociocultural contexts.[60] Researchers and evaluators emphasize the importance of approaches that account for how contextual factors shape experiences, rather than relying on one-size-fits-all methodologies.[61] This

57 Bertram-Troost, "Investigating the impact of religious diversity in schools," 280.

58 Bertram-Troost, "Investigating the impact of religious diversity in schools," 280.

59 Johnny Blair, Ronald F. Czaja, and Edward A. Blair, *Designing surveys: A guide to decisions and procedures*, Third edition ed. (SAGE Publications, Inc., 2014), 261.

60 Garfinkel, *Evaluating interfaith dialogue programs*, 6.

61 Snipes and Correia-Harker, "Implementing Inclusive Interfaith Assessment," 22; Dekker and Meeter, "Evidence-Based Education," 6; Debra J. Rog, "When background becomes foreground: Toward context-sensitive evaluation practice," *New Directions for Evaluation* 2012, no. 135 (2012).

recognition is especially important in interfaith learning, where the societal and behavioral dimensions of learning are influenced by specific contexts.

Several practical recommendations can enhance context sensitivity in the quantitative measurement of interfaith learning outcomes. These include considering contextual factors during data interpretation, mapping out the stakeholders and the contextual factors that influence the implementation of findings into practice, and explicitly integrating context into each aspect of the research process.[62] Our strategy of being sensitive to the context involved incorporating case stories into our measurement tool and situating items within realistic social circumstances rather than in an abstract "vacuum."[63] This approach allowed us to acknowledge inherent power imbalances in interfaith initiatives and recognize the varying perceptions of the religious other depending on the specific identities of this other in different contexts.[64] For instance, the case story above illustrates a situation where a teacher has to make a quick assessment of the individual and societal influences at stake when Leon, a Reformed Christian speaker, leaves the room during Sita's presentation about Indian Hinduism. Instead of asking whether participants would be prepared to listen to others in general, adding such a scenario makes the question of whether, when, and how to listen more realistic.

However, while integrating context into a measurement tool offers benefits in terms of sensitivity and relevance, it also poses challenges. One such challenge is the potential for the tool to become overly context specific. In our study and others in the field of interfaith learning, this challenge manifests itself in the difficulty of applying, generalizing, or comparing outcomes and factors across various contexts.[65] In the example of our case story where Leon, the Christian speaker, walks out while Sita is talking about Indian Hinduism, it is essential to consider that this event is taking place in a European context of Christian/secular privilege and small Hindu

62 Regarding contextual factors, see Dekker and Meeter, "Evidence-Based Education," 6. Regarding implementation, see Taryn Moir, "Why Is Implementation Science Important for Intervention Design and Evaluation Within Educational Settings?," Review, *Frontiers in Education* 3 (2018-July 25-2018). Regarding integration, see Rog, "When background becomes foreground," 26.

63 Janet Finch, "The vignette technique in survey research," *Sociology* 21, no. 1 (1987): 106.

64 cf. Najeeba Syeed, "Interreligious Learning and Intersectionality," in *Asian and Asian American Women in Theology and Religion: Embodying Knowledge*, ed. Pui-lan Kwok (Cham: Palgrave MacMillan, 2020); Stephanie Burrell Storms, "Using Social Justice Vignettes to Prepare Students for Social Action Engagement," *Multicultural Perspectives* 16, no. 1 (2014).

65 Losert, Merkt, and Schweitzer, "In Search of Interreligious Competence," 108.

minorities when deciding how to respond. Therefore, integrating attention to the context comprehensively often requires compromising on the cross-contextual applicability of measurement tools.

Thus, despite the recommendation that any measurement approach should carefully consider the contextual setting of a program to avoid maintaining religious assumptions and biases and to ensure that participants do not respond in a social vacuum, integrating the context in each aspect of the process may require compromising on the cross-contextual applicability of a measurement tool.[66]

2.3. Using Other Methods in Addition to Quantitative Measurement

Integrating other methods alongside quantitative measurement is essential for capturing the complex reality of interfaith learning. As noted by Abu-Nimer in the context of interreligious peacebuilding, "we must not rely exclusively on quantitative data collection or analysis; such data often fails to capture the nuances of peacebuilding work, especially on sensitive change issues."[67] Combining different methods can range from mere comparison to full integration of methods throughout the process.

The choice of a methodological design depends on the specific purpose of using a mixed methods approach. Greene, Caracelli, and Graham identify five such purposes: triangulation, complementarity, development, initiation, and expansion.[68] While all of these purposes can be applied to enhance the evaluation of interfaith learning, we elaborate on the practical implications of three of these purposes that were used in our research project.

In our research project, we used qualitative methods, interviews, and reflection report analysis to develop the case stories and items that made up our quantitative measurement tool, with the intention of making the tool realistic and comprehensive for participants. This approach aligns with the mixed methods purpose of "development," with one method informing the design of another. For example, insights derived from our qualitative data

66 Consider J. Cody Nielsen, Sachi Edwards, and Matthew R. Sayers, "Interfaith?: A Critical Examination of the Interfaith Learning and Development (ILDT) Framework for Religious, Secular, and Spiritual Identities," *Journal of College and Character* 23, no. 4 (2022).

67 Abu-Nimer, "Challenges in Peacebuilding Evaluation," 33.

68 Greene, Caracelli, and Graham, "Toward a Conceptual Framework for Mixed-Method Evaluation Designs."

facilitated the development and refinement of items and questions for our quantitative tool. Another common approach to this is to use quantitative results to identify which participants to invite for follow-up interviews.[69] The practical goal of using mixed methods in this situation is to guide the development of one method, tool, or topic list using findings from another, thereby improving the overall design and validity of the study and making the results more meaningful to participants' lived realities.

A key purpose of using mixed methods, often central when scholars such as Abu-Nimer advocate the use of diverse methods in interfaith learning, is to complement the information that can be provided by quantitative tools alone with narrative descriptions using qualitative methods. The use of complementarity allows one to go beyond surface-level findings and address the underlying factors that influence the data. In our study, quantitative data on participants' leadership skills showed no significant development over time, with participants generally agreeing on all items. At first, we were unsure how to interpret these findings: perhaps the "correct" answers were too obvious, leading to a social desirability bias, or participants' leadership skills may have been cultivated before the program began. By using qualitative methods in addition to quantitative ones, we could identify potential underlying reasons. Our qualitative data suggested that participants did indeed feel that their leadership skills were largely developed before the program began since they had already been working as professionals in the field for a long time.[70] In this way, the results of one method can complement the interpretation of the results of another by providing a different perspective. To achieve this complementarity, different methods should focus on overlapping indicators, such as leadership skills in our example.

Mixed methods can also facilitate the "expansion" of the research scope and breadth of a study.[71] By employing mixed methods to expand a study, different levels of participants' experiences can receive attention, thereby broadening the focus of the study. For example, in our study, the quantitative pilot tool and interviews focused on the outcomes of the Emoena program, while the analysis of reflection reports focused on participants' development and expectations during the program. This approach allowed us to gather information on different aspects of the same topic—that is, participants'

69 Greene, Caracelli, and Graham, "Toward a Conceptual Framework for Mixed-Method Evaluation Designs."

70 Visser et al., "Beyond Interfaith Harmony: How Participants Perceive the Impact of an Interfaith Leadership Program in Flanders and the Netherlands."

71 Greene, Caracelli, and Graham, "Toward a Conceptual Framework for Mixed-Method Evaluation Designs."

development during the program as well as their (perceived) learning outcomes at the end of the program, thus broadening the scope of our understanding of interfaith learning.

Despite these benefits, mixed methods research does pose challenges and it demands a diverse skill set from the researchers: from analyzing the reliability and validity of a measurement tool so as to avoid drawing invalid conclusions, to conducting and coding qualitative interviews while remaining self-reflective. Additionally, mixed methods research requires more time and resources than using a single method, from planning, transcription, and data collection to analysis and integration or comparison.[72] In our pilot study, for example, conducting interviews with participants about their evaluation of the Emoena program took months to prepare, conduct, and analyze, time that may be scarce for researchers and practitioners in the field of interfaith learning. Therefore, while using multiple methods can be an effective way of addressing the story behind the numbers, it requires skills and time that may not always be available to interfaith initiatives. When done correctly and intentionally, the use of both qualitative and quantitative methods can be an effective response to the risk that quantitative measurement does not capture the complexity of interfaith learning.

Conclusion

In this article, we have explored some of the complexities of measuring interfaith learning outcomes, including its underlying purposes, challenges, and implications. While the field of interfaith learning has seen a proliferation of quantitative studies on the outcomes of interfaith initiatives, the questions around why and how to measure them have received little attention. We have argued that there is a risk of misrepresenting or even compromising the complex reality of interfaith learning in the absence of reflection on its underlying purposes and practical implications. This article initiates a conversation about *why* and *how* to measure interfaith learning outcomes.

When done thoughtfully, measuring learning outcomes can provide valuable insights for program improvement; it can foster collaboration by allowing for comparison across initiatives. Quantitative measurement can stimulate self-reflection about what we do and why we think it matters.

72 John W. Creswell and Vicki L. Plano Clark, *Designing and conducting mixed methods research* (Thousand Oaks, California: SAGE Publications, 2007), 181.

However, it can also oversimplify the work of practitioners and reduce it to metrics for funding or performance reporting. Such oversimplification can then have a negative impact and shift the focus of what is considered important. Therefore, we should carefully consider why and when measurement is helpful and how to prevent it from overshadowing the underlying ideals and content of interfaith learning.[73]

Any attempt to measure the outcomes of interfaith learning should recognize that quantitative measurement can provide helpful insights but cannot capture all aspects of the messy, complex, and multifaceted nature of interfaith learning. We have proposed three considerations for measuring interfaith learning outcomes: involving multiple stakeholders, remaining sensitive to the context, and using other methods in addition to quantitative ones. These suggestions stimulate reflection on finding an appropriate balance between more resource-intensive, dialogical processes that involve multiple stakeholders and integrate mixed methods at different stages of the process, and less intensive and more outcome-oriented approaches. It is crucial to avoid reducing the richness of interfaith learning to mere numbers and to appreciate its unmeasurable aspects.

This article has brought together insights from the educational sciences, peacebuilding studies, and other disciplines in order to stimulate reflection and prompt discussion about the (im)possibility of measuring interfaith learning outcomes among scholars and practitioners of interfaith learning. The dynamics related to measurement that we have described here are not unique to interfaith initiatives, but our field can be seen as a good example of the challenges involved because of its inherently contextual, intangible, and socio-political nature. Our response to these dynamics can also be used for other disciplines. By initiating a conversation about the challenges, assumptions, and implications of measuring interfaith learning outcomes, we aim to promote unassuming, realistic, and nuanced approaches to understanding the outcomes of interfaith learning.

Acknowledgments
The authors thank Ineke Nagel, assistant professor of Sociology at VU Amsterdam, for her assistance in analyzing and interpreting the data from our quantitative pilot study.

73 Consider Biesta, *Good Education in an Age of Measurement*.

Hannah J. Visser *is a postdoctoral researcher and facilitator of the Emoena interfaith leadership program at the School of Religion and Theology, Vrije Universiteit Amsterdam. She is also a postdoctoral researcher at the Research Institute for Child Development and Education, University of Amsterdam. Her research focuses on religion, diversity and dialogue in educational settings. She obtained her PhD in January 2025 with a dissertation on the impact of interfaith learning.*

Gerdien D. Bertram-Troost *is full professor of Religious Education at the School of Religion and Theology at the Vrije Universiteit Amsterdam. Her main research interests lie in the fields of education, developmental psychology, theology, and their overlap.*

Marianne Moyaert *is full professor of Comparative Theology of Religions at the Research Unit Systematic Theology and the Study of Religions, Faculty of Theology and Religious Studies at the KU Leuwen. She specializes in the (comparative) theology of religions, interreligious hermeneutics, and Jewish-Christian relations*

Anke I. Liefbroer *is endowed professor of Interconnection Psychiatry and Theology and associate professor of Psychology of Religion and Spiritual Care at Tilburg School of Catholic Theology. Her main research interests lie in the fields of (interfaith) spiritual care and psychology of religion.*