THE DEVELOPMENT AND VALIDATION OF AN INTERCULTURAL COMPETENCIES ASSESSMENT INSTRUMENT FOR K-12 IN-SERVICE EDUCATORS

by

David Ellsworth Lynn

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Educational Research, Measurement, and Evaluation

Charlotte

2023

Approved by:

_____
Dr. Jae Hoon Lim

_____
Dr. Stella Kim

_____
Dr. Richard G. Lambert

_____
Dr. Erik Jon Byker

ABSTRACT

DAVID ELLSWORTH LYNN. The Development and Validation of an Intercultural Competencies Assessment Instrument for K-12 Educators
(Under the direction of DR. JAE HOON LIM and DR. STELLA KIM)

As schools adapt curriculum and learning environments to better prepare students for entry into an increasingly globalized society, cultivating intercultural competencies in K-12 in-service educators is of heightened importance. The purpose of this study was to develop and validate a new instrument designed to assess these competencies called the Intercultural Competency Measure for Educators (ICME). Byram (1997) defines intercultural competencies as the ability to effectively communicate, understand, and work with people from diverse cultural backgrounds. Deardorff (2006) adds to this a call for action, which lends itself to the critical cosmopolitanism framework that guides this study. A pilot study was used to develop a four-factor theoretical intercultural competencies framework through a process defined in this study. Reliability and validity were examined using data collected from K-12 in-service educators at schools in the United States and Canada. An Exploratory Factor Analysis suggested a revision of the constructs to include five factors: Curriculum, Diverse Student Inclusion, Cross-Cultural Openness, Collaboration and Adaptation, and Systematic Awareness. Construct validity was tested using Confirmatory Factor Analysis and supported by examining demographic data using parametric tests. The emergence of a factor related to systematic awareness highlights teachers' increased role in addressing the root causes of inequity in schools. The five-factor model provides a framework for schools wishing to further develop and assess intercultural competencies growth in teachers.

DEDICATION

With Love to My Family

In memory of my maternal grandparents - Hinda K. Nathan & Kurt C. Nathan, Ph.D.
for exemplifying intercultural competencies, modeling lifelong learning,
and nurturing my wanderlust

To my mother and stepfather - Vivian N. Campbell & Ronald J. Campbell
for modeling lives devoted to service and encouraging me each step of the way

To my father - Gordon D. Lynn
for modeling hard work, precision, and independence

To my mother-in-law and father-in-law –Zheng Qian Zhao & Jian Guo Qi
for modeling courage, wisdom, and sacrifice

To my children - Drake, E'Ara, Margaret, and Nathan
for bringing joy, needed distractions, a few grey hairs, and endless hope for the future

To my partner in the journey of life – Ming Qi, MD
for humbling me with her beauty, knowledge, and heroism
for making me feel like the luckiest person alive when I am in her presence

ACKNOWLEDGEMENTS

My maternal grandfather, Kurt C. Nathan, waited until his children were in college before commencing his post-secondary education. His dreams were put on hold after fleeing Germany during the Holocaust, serving in the US Armed Forces during World War II, and raising a family with my grandmother, Hinda. With steady perseverance, he earned his Ph.D. in May 1973 at the age of 57. Fifty years later, I am honored to follow in his footsteps and grateful for his influence.

My deepest appreciation to my advisor and Committee Chair, Dr. Jae Hoon Lim, and Co-Chair, Dr. Stella Kim, for setting high standards, expanding my perspective, and offering wise counsel. Thank you to Dr. Richard Lambert, a member of my committee, who nurtured me during my infancy in statistics and mentored me during two research internships. I am fortunate to have Dr. Erik Byker, whose research into critical cosmopolitanism is central to this study, as a committee member. Thank you to the UNC-Charlotte faculty who have patiently supported my growth while challenging me over the past five years: Dr. Sandra Dika, Dr. Claudia Flowers, Dr. Florence Martin, Dr. Xiaoxia Newton, Dr. Carl Westin, and. Dr. Greg Wiggan.

I am full of gratitude and admiration for my ERME peers who offered incredible comradery and ensured good times along this research journey. First and foremost, I am indebted to Dr. Bryndle Bottoms for being an exceptional research partner, sounding board, and motivational support. Thank you to the outstanding project partners and idea generators in our cohort: Dr. Scott Holcomb. Dr. Courtney Green, Zhi Li, Kirstin Villanueva, and Yi Wang. Thank you to Youlia Tzenova for collaboration during my first two years.

Thank you to my colleagues at Charlotte Country Day School for ongoing support and insight over the past five years: Tanya Andrysiak, Tianna Butler, Angelina Ferrari Segovia, Joe Gardner, Marcel Gauthier, Nancy Hackett, Matt Less, Bill Mulcahy, Jessica Newman, Kato

Nims, Rebecca Patterson, Mark Reed, Kathy Rodgers, Warren Sepkowitz, Kasey Short, Scott Waybright, and Brian Wise; to colleagues linked to my years at the American School of Paris, Mark Ulfers, Aaron Hubbard, and Michael Kopp, for their global perspective; to Dr. Joseph Hoff of the UNC Charlotte Office of International Programs for research support; to the team at Charlotte's Mayor's Youth Employment Program (MYEP) for partnership during my ERME internship project: Andrew Knoblich, Sterling Oliver, Bethany Rodriguez, and Dawn Hill; to Ljubomir Stambuk and Charlotte Klopp of the World Affairs Council of Charlotte; and to fellow Board members of the North Carolina Association of Research in Education (NCARE).

Data collected in this study were made possible thanks to collaboration with the Global Education Benchmark Group (GEBG), the support of GEBG Executive Director Dr. Clare Sisisky, and regular discussions with Dr. Kassandra Brenot and Dr. Aric Vissar. Thank you to the global education leaders and faculty at the 16 GEBG schools who participated in this study.

The loving patience and unyielding support of my family was instrumental each step of the way. Thank you to my mother, Vivian Campbell, stepfather, Ronald Campbell, and in-laws, Zheng Qian Zhao and Jian Guo Qi, for always believing in me. To Drake, for Excel brilliance and technical support; to E'Ara, for motivational sticky notes on my computer; to Maggie, who ensured evening story time was prioritized over journal reviews; and to Nathan, who was not yet born when I was accepted into the ERME program, for regular games of UNO.

Finally, with eternal love and appreciation, to Ming, who greenlighted this journey back into academia. She sacrificed more than anyone to ensure I could get to this point. When the pandemic brought unimaginable personal and professional challenges, she was always ready to lean in to offer more. Ming, while I am proud of you for being recognized as Charlotte's 2022 ER Physician of the Year, you've been and always will be number one in my book.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Globalization has significantly elevated the recognition of intercultural competencies as an educational priority. The United Nations Educational, Scientific and Cultural Organization (UNESCO) and the Organization for Economic Co-operation and Development (OECD), along with the United States Department of Education and Council of Europe, have all placed a heightened emphasis on global education as a means of preparing students for an increasingly interconnected world (Martorana et al. 2021; OECD, 2018). This trend responds to the expansion of human migration, changes in how people work and communicate, awareness of sustainable development priorities, and the evolving social patterns that have further promoted cross-cultural interactions (Kerkhoff, 2017). Government and global organizations are under increased pressure to ensure young learners cultivate the attitudes, knowledge, and skills necessary to navigate the complexities of globalization while remaining adaptable to ongoing changes (Deardorff, 2011).

Increased global interconnectedness has raised questions regarding the overall purpose of global education. Two major perspectives that have influenced the evolving conversation of global education are neoliberalism and critical cosmopolitanism (Byker, 2016). While neoliberalism has offered a practical rationale, critical cosmopolitanism provides a framework that responds to contemporary social, humanitarian, and political concerns. Neoliberalism is embedded in a capitalistic perspective and emphasizes the need for workers prepared to compete in the global marketplace with the skills and connections that provide strategic advantages (Schultz, 2007; Goren & Yemini, 2017, Kerkhoff, 2022). The desire to inspire “citizens of the world” defines cosmopolitanism, guided by the promotion of intercultural dialogue, respect for diversity, openness toward differing perspectives, and a capacity for critical self-reflection regarding one’s own culture and worldview (Byker, 2016; Oxley & Morris, 2013). Critical

cosmopolitanism recognizes global interconnectedness, highlights inequalities, and inspires change through active observation, listening, empathizing, and critical self-reflection (Hauserwas et al., 2021). Cosmopolitanism aligns with intercultural competencies related to attitudes, knowledge, and skills, while also inspiring a desire to take action.

Schools have steadily transitioned from primarily fostering nationalistic tendencies to an increased embrace of the development of intercultural competencies across the curriculum and at all grade levels (Byram, 2021; OECD, 2018). In the United States, this is further motivated by the increased diversity of students in American classrooms. Between 2009 and 2018, the percentage of public school students who identified as people of color increased from 47% to 53%, with students of Hispanic heritage rising from 22% to 27% (National Center for Education Statistics, 2021). The development of intercultural competencies aligns with an increased emphasis on diversity awareness by helping to broaden perspectives and promoting an understanding of inequalities and power dynamics (Bernardes et al., 2019; Choi & Shin, 2016). While knowledge of cultural facts, current events, and language fluency are important, it is not enough to address the responsibility schools must prepare global citizens (Deardorff, 2016). Intercultural competencies development promotes interpersonal growth, including an awareness of others, empathy, and adaptability (Biautti, 2018; Morais & Ogden, 2011). This further encourages students to suspend judgment when a cross-cultural conflict emerges and proactively seek new perspectives (Biautti, 2018). While these competencies are important in a school environment, they are broadly applicable and relevant in contemporary society.

While society has embraced the development of intercultural competencies in students, the development of similar attitudes, knowledge, and skills in educators has been given less attention by researchers and school leaders. Educators tasked with designing and delivering

intercultural competencies instruction may not have been exposed to related themes and perspectives when they were students themselves or in teacher training (Hauerwas & Kerkhoff, 2021; Scheicher, 2015). In addition, educators delivering intercultural instruction often face challenges in navigating increasingly diverse classrooms while addressing inequities and power structures within the school environment (Deardorff, 2006; Larsen & Searle, 2017). Often, educators are confronted with unclear definitions of intercultural competencies, vague direction when designing a globally themed curriculum, and limited resources for creating intercultural learning environments (Colvin, & Edwards, 2018; Smolcic & Katunich, 2017). The insufficient guidance on how to assess the intercultural competencies of K-12 educators has handicapped the ability of school leaders to identify best-practice or select targeted approaches to professional development. This has also led to inconsistency concerning how intercultural competencies are taught from classroom to classroom and school to school.

Statement of the Problem

The intercultural competencies modeled by educators impact the effectiveness of their ability to design and implement related themes in learning environments. Intercultural competencies skills equip teachers with the attitudes, knowledge, and skills to contribute to actions that promote the inclusion of culturally diverse learners and ensure all students have exposure to broad global perspectives. Implementation in schools is complicated by the challenge of establishing a common definition for intercultural competencies (Deardorff, 2006). This lack of a clear framework has inhibited school leaders from adequately assessing intercultural competencies in educators and providing professional development opportunities that lead to meaningful growth (Kerkhoff, 2017).

Research into the development and assessment of intercultural competencies in K-12 educators is primarily limited to studies involving pre-service teachers who are easily accessible to university researchers (Kerkhoff, 2022). Due to the complexity of quantifying intercultural competencies, most studies have relied on a qualitative approach to assessment based on interviews, observations, group discussions, and interviews. Byram (2021) notes that even with the shortcomings of a quantifiable assessment, the development of a psychometrically sound scale would likely be the most beneficial means to serve the needs of schools. In addition to providing insight into individual educator competencies, a scale providing quantitative data could offer a clearer perspective of the growth of intercultural competencies throughout a school community.

Current intercultural competencies scale development provides only limited insight regarding applicability to educational settings. The Intercultural Development Inventory (IDI), based on 50 multiple-choice items, demonstrates high cross-cultural validity and reliability when measuring competencies themes (Álvarez Valdivia & González Montoto, 2018; Hammer, 2011). The IDI places participants on a seven-part continuum that categorizes the individual's level of intercultural competencies development. However, the IDI was not developed with educational settings in mind and does not consider the specific competencies that enable educators to navigate a globally diverse learning environment (He et al. 2017). While the IDI provides insight into a participant's personal perspective, it does not directly promote critical reflections concerning curriculum design, student engagement, or classroom design.

Recent studies examine assessment instruments designed for educational settings but are limited in applicability to the entire K-12 learning environment. Kerkhoff's (2017) Teaching for Global Readiness (TGR) scale provides an assessment instrument specifically designed to apply

to educators. Instead of placing the participants on a continuum like the IDI, the TGR offers insight into the application of intercultural competencies to instructional practice. However, the TGR is not generally applicable to K-12 education and focuses heavily on literature curriculum and the use of technology. An intercultural assessment instrument released by the OECD called the Program for International Student Awareness (PISA) has attempted to establish worldwide benchmarks for global education but has been more heavily focused on finding common global priorities, including a focus on environmental sustainability, rather than examining how to embrace and navigate cultural differences (Byram, 2021). Despite the critical limitations of the three scales in assessing educators' intercultural competencies, the scales and related research literature provide valuable insight into which domains and factors would be most relevant when measuring teachers' intercultural competencies throughout the general school setting.

A theory-based and empirically validated K-12 educators' intercultural competencies assessment instrument will enable teachers to understand their intercultural competencies better while critically reflecting on their global pedagogical practices (Hur & Sur, 2018). This scale will also provide school leaders with needed data to guide their faculty's professional development, assess interventions through pre- and post-data collection and analysis, and chart longitudinal growth (Okken et al., 2019). This assessment instrument will need to be examined for content validity in its creation and construct validity to ensure the data collected are consistent with its intended use. A validated assessment instrument has the potential to provide insight into key attributes and effective interventions that best serve to promote intercultural competencies. It will also help provide opportunities for professional reflection and inspiration on how global curriculum is developed, how global learning environments are designed, how

students of all cultural backgrounds are engaged, and how the personal development of an educator can impact their ability to model the competencies they desire to teach.

A proposed K-12 educator intercultural competencies scale, called the Intercultural Competency Measure for Educators (ICME) was examined for content and construct validity based on an original four-factor theoretical model. The dependent variables were theoretical factors determined based on a review of the literature, pilot study, experts' input, and practical experience of the researcher in the field of global education. In the original theoretical four-factor model, three of the four factors directly related to educational settings and apply to all grade levels and content areas: (a) global curriculum content, (b) intercultural learning environment, and (c) student engagement. A fourth factor, personal intercultural competencies growth, focused on how educators' model intercultural competencies and extend them beyond the school environment. The ICME was designed to not only assess educators' intercultural competencies, but also to spark critical reflection and guide personal and professional development. Despite the crucial role of intercultural competencies in global education, scholars have yet to give sufficient attention toward creating a psychometrically valid scale that offers robust and reliable quantitative data on educators' intercultural competencies, particularly in ways that are applicable across the curriculum and throughout the entire K-12 learning environment. Therefore, the ICME would fill a significant void in the existing literature by offering a psychometrically sound scale well-grounded in the current directions of global education and intercultural competencies.

Purpose of the Study

The purpose of this study was to develop and validate the ICME utilizing data collected from an in-service faculty assessment instrument that was distributed to schools in the United

States and Canada. The initial version of the ICME explored the four domains of intercultural competencies: (a) Attitudes, (b) Knowledge, (c) Skills, and (d) Action as they related to the four-factor theoretical model, (a) global curriculum content, (b) intercultural learning environment, (c) student engagement, and (d) personal intercultural competencies growth. Each of the four theoretical factors of intercultural competencies fully crosses with the four domains. The domain of *Attitude* relates to respect, openness, curiosity, and tolerance (Deardorff, 2011). The domain of *Skills* includes adapting behaviors and mannerisms, analyzing the cultural context, and utilizing the abilities to listen, observe, evaluate, interpret, and reflect (Deardroff, 2011). The domain of *Knowledge* reflects insight into history, culture, language, traditions, and perspectives in ways that often promote an awareness of diversity, inequities, and power structures (Martorana et al., 2021). The domain of *Action* reflects the cosmopolitism outcome of critical reflection contributing to a desire to impact change either locally or globally. The theoretical factor *global curriculum content* includes how classroom lessons and academic content reflected global and cultural themes. The theoretical factor *intercultural learning environment* reflects how the classroom and overall school setting imitated cultural inclusivity and global connections. The theoretical factor *student engagement* relates to how globally diverse students' learning needs, perspectives, voices, and contributions are recognized. The final theoretical factor *personal intercultural competencies growth* includes how participants explore, adapt to, and reflect upon new cultural experiences and engage with people of diverse cultural backgrounds beyond school and teaching context.

This study also conducts an analysis to examine how findings from the four-factor theoretical model were predicted by demographic characteristics traditionally associated with contributing to intercultural competencies. These characteristics included: (a) core teaching

content area; (b) proficiency in a second language, (c) length of time lived outside of the country of current primary citizenship, and (d) race/ethnicity, (e) the number of years in the teaching profession.

Research Questions

This study examined the design of the ICME while seeking to validate a four-factor theoretical model based on data collected from K-12 educators at independent schools. This study aimed to provide additional validity evidence of the ICME through examination of the relationship between specific demographic characteristics traditionally associated with contributing to intercultural competencies. Specifically, the following three research questions guided the study.

1. *What is the extent to which content validity is established in the design of the Intercultural Competency Measure for Educators (ICME)?*
2. *To what extent does the ICME reveal construct validity evidence for its intended use?*
3. *To what extent do demographic characteristics commonly associated with contributing to intercultural competencies growth predict results of the ICME based on the four-factor theoretical model?*

Theoretical Framework

The guiding framework for this study is derived from Byram's (1997) research-based constructs, which he articulated as "intercultural communicative competencies (p. 36)." Byram's study initially focused on enhancing the teaching of modern world languages by emphasizing developing competencies related to knowledge of self and others, global attitudes, the ability to interpret and connect, and the capacity for critical cultural awareness. The framework for intercultural competencies was later broadened to include all subject areas and teaching levels,

with an emphasis on (a) attitudes, (b) knowledge, and (c) skills (Byram, 2021; Deardorff, 2006; Kerkhoff, 2017). This study blends the framework for intercultural competencies with critical cosmopolitanism to add emphasis on action. Critical Cosmopolitanism calls for a heightened level of consciousness guided by a reflective process that leads to an enhanced awareness of political, moral, cultural, and economic relationships. This recognition leads to critical thought, providing a path toward desired action and change (Byker, 2016). Critical cosmopolitanism recognizes the responsibility teachers have to construct relevant global curriculum, engage diverse students, and respond to inequities and power relationships in the learning community.

Research Design

This quantitative study began with the development of the ICME based on the domains of intercultural competencies and critical cosmopolitanism. A pilot study examining data from 102 survey respondents used an exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) to create a four-factor theoretical model for K-12 educator intercultural competencies growth. The results of the EFA and CFA initially recommended three factors. Two theoretical factors, building a global curriculum and personal global competencies, were retained for this study. The third factor, student competencies growth, was later split into two factors, global learning environment and student engagement, based on professional insight, ongoing review of the literature, and an examination of items that did not fit the original pilot factor model.

A 30-item version of the ICME, based on the four-factor theoretical model, was distributed to educators at independent schools with membership in the Global Education Benchmark Group (GEBG), a network of schools in the United States and abroad that are linked by a shared desire to promote global learning. In addition to the assessment items, participants were asked to provide to provide information about nine personal demographic variables and

three school-related demographic variables. The data collection period took place between May and June 2022.

Once the data were collected from an adequate sample size, descriptive statistics were examined prior to checking for missing data and multivariate outliers. An EFA was used to examine if the results matched the theoretical model. A CFA was used to examine the IDIAM's construct validity and reliability. An examination of goodness of fit was used, along with feedback from a panel of experts, to explore the plausibility of the model. Recommendations from the modification indices and panel of experts were used to examine localized fit and make changes.

A finalized model was examined using parametric tests including independent sample's t-test and one-way ANOVA to check assumptions and examine how demographic data, specifically independent variables traditionally associated with contributing to intercultural competencies growth, relate to dependent variables found in the ICME. These findings helped to provide insight into both the construct validity and discriminant validity of the model.

Significance of the Study

As more schools embrace intercultural competencies as an educational priority, the ability to clearly define and assess these competencies in educators is of increased importance. While a growing number of research studies examine intercultural competencies development in pre-service educators, studies focusing on in-service educators remain limited. Previous studies of in-service educators mostly rely on qualitative data collected from a limited sample size. Prior intercultural competencies assessment instruments used in research were either designed for broader audiences, not addressing the needs of educators, or were targeting specific educational niches and not applicable for cross-curricular use. This study utilizes quantitative data from a

theory-driven assessment instrument designed specifically for in-service K-12 educators regardless of subject area or grade level. The ICME promotes critical reflection in educators and guides professional development that will enhance how intercultural competencies are further developed and embraced.

The validation of the ICME will promote the use of this instrument to gather quantitative data that can help both an individual educator and school-wide assessments (Kerkhoff, 2017). The demographic data collected through this study will help examine whether specific participant characteristics or lived experiences contribute to the growth of intercultural competencies. This data can potentially guide the development, funding, and evaluation of professional development interventions (Morais & Ogden, 2011). Opportunities for longitudinal research will open as parallel forms of the assessment scale are created, allowing for pre- and post-assessments. This will allow for a broader understanding of the effect of targeted interventions.

Delimitations

This study includes data collected from educators currently teaching at K-12 independent schools. The selection criteria for participating schools included being K-12 in focus, located in the United States or Canada, and members of the Global Education Benchmark Group (GEBG) network. Data were collected using Qualtrics through the IDIAM alongside a demographic questionnaire.

Definitions of Terms

Actions. A desire to apply one's attitudes, knowledge, and skills to promote change either locally or globally.

Attitudes. This refers to behaviors related to respect, openness, curiosity, and tolerance (Deardorff, 2011).

Construct Validity. The extent to which the items measured accurately assesses the hypothetical constructs (Huck, 2004, p. 95)

Content Validity. The extent to which the items collectively cover the material intended to be measured by the instrument (Huck, 2004, p. 89)

Cosmopolitanism. Striving to become a "Citizen of the world" through the recognition of global diversity, respect for differing perspectives, desire to interact cross-culturally, capacity for critical reflection, and willingness to pursue action and change (Byker, 2016).

Intercultural Competencies. The ability to navigate diverse cultures through broadening knowledge of oneself and others; developing the skills to interpret and relate; cultivating the desire to discover and interact; valuing the belief and behaviors of diverse groups; and relativizing oneself (Byram, 2021).

Knowledge. Evidence of insight into history, culture, language, traditions, and perspectives in ways that often promote an awareness of diversity, inequities, and power structures (Martorana et al., 2021).

Neoliberalism. Seeking global exposure as a means to an economic advantage while providing entry into an environment of market competition based on access to power and privilege (Pashby et al., 2020).

Reliability. The consistency of scale scores when replicated using similar assessment procedures (AERA, APA, & NCME, 2014, p. 33).

Self-Reporting Bias. When an instrument relies on self-reported data from a participant, there is a risk of bias and construct-irrelevant variance (AERA, APA, & NCME, 2014, p. 56).

Skills. The ability to adapt behaviors and mannerisms while analyzing the cultural context, while utilizing the abilities to listen, observe, evaluate, interpret, and reflect (Deardroff, 2011).

Validity. The extent to which the data and theory support the interpretation of scale scores for the intended use of the assessment instrument (AERA, APA, & NCME, 2014, p. 11).

Dissertation Organization

This dissertation is organized in five chapters, followed by references and appendices. Chapter Two reviews the literature related to intercultural competencies, neoliberalism, cosmopolitanism, and past examples of assessment instruments. Chapter Three provides detailed descriptions about the methodology of the study. This chapter explains how the ICME was designed, the procedures for collecting the data, and how the data were analyzed. Chapter Four provides an overview of the results and findings. Chapter Five offers a summary of key findings followed by scholarly discussion, practical and policy implications, limitations, and recommendations.

Summary

Martorana et al. (2021) note that educators are the key social actors responsible for helping forge global citizens ready to take on the challenges of an increasingly interconnected society. Educators need opportunities for further reflection and professional development related to intercultural competencies growth (Colvin & Edwards, 2018). A theory-driven and

empirically validated formative assessment can promote guided development and chronicle longitudinal growth (Hammer, 2012). The data collected through a theoretically and psychometrically sound scale have the potential to measure the effect of target interventions and provide insight into growth at both the individual educator and school-wide levels.

As global education becomes a priority in K-12 schools, the need for a validated theoretical model will provide direction into global curriculum development, cross-cultural student engagement, learning environment design, and personal intercultural competencies growth. Data from a validated scale allows for insight into which demographic variables contribute to intercultural competencies development and prompts new approaches towards professional development and other interventions to be considered. This data provides teachers with a needed intercultural competencies framework while promoting opportunities for critical reflection and targeted action.

Research into the nurturing of intercultural competencies in school environments remains limited. Few assessment instruments exist to measure intercultural competencies development in educators, making it hard to assess the concurrent validity of the ICME. This research study aimed to contribute to this growing body of knowledge on educator intercultural competencies and provide validation data that will be useful on a practical level to schools while helping to fuel future research investigations.

CHAPTER 2: LITERATURE REVIEW

Our increasingly globally diverse society highlights the importance of intercultural competencies related to knowledge of self and others, global attitudes, the ability to interpret and connect, and the capacity for critical cultural awareness (Byker & Putnam, 2019; Byram, 2021; Deardorff, 2011; Smolcic & Katunich, 2017). Over the past decade, the United States Department of Education (USDOE) and the United Nations Educational, Scientific and Cultural Organization (UNESCO) have responded with an emphasis on the importance of schools in preparing students for an ever-increasingly global society (Kerkhoff, 2017; Zhao, 2010). The Organization for Economic Co-operation and Development (OECD) has highlighted the need for change in approaches to teaching to ensure students develop the competencies necessary for navigating an ever-changing world (Colvin & Edwards, 2018; Hauerwas & Kerkhoff, 2021; OECD, 2018; Scheicher, 2015). On a practical level, educators have had to navigate challenges based on global diversity and language differences in the classroom that have led to inequities and impediments to learning (Deardorff, 2006; Kerkhoff & Cloud, 2020; Larsen & Searle, 2017). Smolcic and Katunich (2017) recognized that the development of teacher intercultural competencies is full of complexities and with no simplistic solutions.

This literature review will examine efforts by researchers to define intercultural competencies and and their relevancy to educators and explore assessment strategies that seek to measure these competencies. Developing a theoretically grounded scale and measurement instrument for assessing intercultural competencies will be critical as schools devote more time and financial resources toward offering related professional development opportunities (Fantini, 2009; Kerkhoff, 2017; Morais & Ogden, 2011). Professional development experiences abroad for in-service educators is one example of an intervention that could potentially benefit from the

insight provided by an assessment instrument (Baecher & Chung, 2020; He et al., 2017; Kerkhoff, 2017; Nielsen et al., 2019; Okken et al., 2019). In seeking to address the challenge of designing intercultural competencies measurements for educators, it is first essential to explore the multiple views on the purpose of global education and the prior research that defines which intercultural competencies best relate to school settings.

Neoliberal and Cosmopolitanism Views on Global Education

Identifying the essential intercultural competencies for teachers remains as elusive as the diverse viewpoints regarding global education from which they originate. On a practical level, the neoliberal perspective on global education emphasizes that schools prepare students for entry into an environment of market competition to serve a human capital function and maximize performance in the workforce (Byker & Putman, 2019; Pashby et al., 2020). Byker (2016) notes the exclusionary and sometimes exploitative policies that occur as a result of a neoliberal outlook. Goren and Yemini (2017) highlight how access to global education provides a strategic economic advantage to those with privileged schooling while leaving out lower-income students. Bernardes et al. (2019) observed that 46 Canadian teacher candidates in teaching practicums in Kenya had difficulty breaking away from a neoliberal global perspective and noted the reinforcement of pre-existing perceptions of power relationships. Researchers have identified a need to break away from a neoliberal focus of global education that prioritizes emerging markets and competition, with an emphasis on attitudes and perceptions and an awareness of inequalities and power dynamics (Bernardes et al., 2019; Choi & Shin, 2016; O'Connor & Zeichner, 2011). This shift away from the neoliberal viewpoint fueled a desire to emphasize awareness and actions by embracing the virtues of global citizenship through a cosmopolitan perspective.

Cosmopolitanism, which translates as “citizen of the world,” promotes intercultural dialogue and understanding through recognition of diversity, respect for different perspectives, and a capacity for critical reflection on their own culture and worldview (Byker, 2016; Kerkhoff, 2017; Müller et al., 2020; Oxley & Morris, 2013; Pashby et al., 2020; van Werven et al., 2023; Wahlström, 2014). In educational settings, Critical Cosmopolitan Theory places a humanitarian emphasis on the development of a critically conscious global citizen (Byker, 2016; Byker & Ezelle-Thomas, 2021). Oxley and Morris (2013) note the range of perceptions of global citizenship and highlight the incorporation of cosmopolitanism into their topography of crucial approaches. The researchers drew critical links between cosmopolitan approaches to global education and enhancing political, moral, cultural, and economic relationships. In addition to simply breaking down nationalistic outlooks, critical cosmopolitanism extends beyond acts of tolerance by embracing the interdependence of the world through actively observing, listening, and understanding different perspectives (Hauserwas et al., 2021; Kerkhoff, 2017; Wahlström, 2014). The process of expanding one’s level of consciousness has the potential to transition a learner from simply “reading the world” to “rewriting the world” (Byker & Putman, 2019, p. 88). As a result, Critical Cosmopolitanism emphasizes the importance of social responsibility, focusing less on one’s identity in the world and more on one’s readiness for action to change it (Byker & Ezelle-Thomas, 2021; Kerhoff, 2017; Oxley & Morris, 2013). Educators who embrace cosmopolitanism must be prepared to distinguish between the desire to broaden their understanding of diverse perspectives and adopting a more critical approach that addresses global inequalities with a desire to actively pursue change.

The Emergence of Global Education

The competing viewpoints of neoliberalism and cosmopolitanism highlight the complexity of incorporating a unified approach toward addressing global learning in educational settings. There is currently no universal definition for global education that is agreed upon by researchers (Deardorff, 2006; Guo, 2014; Pashby et al., 2020; Tichnor-Wagner et al., 2016; van Werven et al., 2023). The lack of a unified construct definition presents challenges to educators who aim to promote the development of global perspectives in the classroom. Van Werven et al. (2023) utilized the Delphi method to develop a global citizenship typology based on input from educators familiar with global competencies. Researchers surmised that cultural curiosity, awareness of interconnectedness, valuing cultural differences while recognizing cultural barriers, and clear communication were essential elements. Other researchers note the importance of infusing globally themed content across the curriculum, including current world events (Byram, 2021, p. 122; Colvin & Edwards, 2018; Ghosn, 2020; Guo, 2014; Kerkhoff, 2017; Kerkhoff & Cloud, 2020; Putman & Byker, 2020; Tichnor-Wagner et al., 2016; van Werven et al., 2023). Providing time for reflection, particularly critical examination of one's worldview, is also vital in the transformational development of global learners (Guo, 2014; Okken et al., 2019; Parkhouse et al., 2016; Patterson, 2015). Targeted teacher training and professional development are needed to ensure educators have the skills necessary to infuse global learning in their classrooms.

Nurturing a capacity for critical cultural reflection is an essential element of cultivating intercultural competencies in educators (Byram, 2021, p.140; Ghosn, 2020; He et al., 2017; Hur & Suh, 2018; Kerkhoff & Cloud, 2020; Okken et al.; 2019; Trilokekar & Kukar, 2011). Byram (2021, p.140) emphasizes the importance of critically evaluating one's own culture alongside developing an understanding of an outside culture. Through careful reflection, one can promote

self-understanding while separating from one's ethnocentrism and prior assumptions. Positive cross-cultural interactions are promoted due to broadened critical perspectives (Bernardes et al., 2019; Kambutu & Nganga, 2008; Kerkhoff & Cloud, 2020; Martorana et al., 2021; Portera, 2014). Consciousness of one's lack of understanding can promote awareness of the privileges many educators possess as members of a dominant culture, fostering recognition of ways in which normalized social hierarchies impact power structures and relationships (Byram, 2021; Chio & Shin, 2016; Deardorf, 2011; Muller et at., 2020; Trilokekar & Kukar, 2011). As educational settings a adapted to be more welcoming and culturally inclusive, this heightened sense of critical self-awareness and insight into the circumstances of others is essential. Deardorff (2011) notes that critical self-reflection and global awareness are a continuous and never-ending process requiring opportunities for purposeful reflection.

Critical cultural awareness and reflexivity promote an educator's capacity for empathy and the ability to value multiple perspectives (Guo, 2014; Hauerwas et al., 2021; Kambutu & Nganga, 2008; Kerkhoff, 2017; Kerkhoff & Cloud, 2020; O'Connor & Zeichner, 2011; Tichnor-Wagner et al., 2016; Trilokekar & Kukar, 2011). Empathy promotes respect for students' ideas, beliefs, and viewpoints, even when contrasting with those of the dominant culture (Deardorff, 2006; Deardorff, 2011; Martorana et al., 2021; Zhao, 2010). Guo (2014) notes that educators who wish to prepare globally responsible citizens must first respond to social and cultural norms in the classroom, including attention to diverse perspectives and multicultural traditions, before addressing the content needs of a global curriculum. Putman and Byker (2020) argue that traits related to empathy and respect for differences cannot be fully developed without curriculum content reflective of the diverse cultures and students' lived experiences represented in the classroom. The development of teacher empathy and respect for the diversity of student cultures

represented and the delivery of a rich global curriculum are linked and mutually connected to the success of the other.

An increased capacity for empathy mixed with an appreciation for the intersections of diverse cultures and heightened awareness of local and global issues can result in a desire to challenge the status quo and pursue actions that lead to change (Larsen & Searle; 2017; Martorana et al., 2021; Morais & Ogden, 2011; Trilokekar & Kukar, 2011). Hauserwas et al. (2021) connect the development of critically reflective and globally competent individuals with the increased motivation to better understand and take action when confronted with inequality and oppression at the local or global level. Educators have the opportunity to take immediate action by collaborating with professional colleagues on addressing the integration of global curriculum and supporting the learning and emotional needs of all students (Ghosn, 2020; He et al.; 2017; Kerkhoff & Cloud, 2020; van Werven et al., 2023). Increased critical cultural awareness can also lead to a desire for transformational change that addresses systemic inequities related to human rights, power, privilege, and oppression in society (Kerkhoff & Cloud, 2020; Larsen & Searle; 2017; Martorana et al., 2021; Morais & Ogden, 2011; O'Connor & Zeichner, 2011; Zhao et al., 2009). A deeper understanding of intercultural competencies provides a foundation for educators and school leaders wishing to model thoughtful and purposeful responses to increasingly globally diverse classrooms in ways that inspire students as they prepare for civic participation in addressing the complexities of an ever-evolving and interconnected society.

Defining Intercultural Competencies

This literature review builds on the framework presented by Byram (1997), which introduces the concept of intercultural communicative competencies, primarily focused on

language educators, as a means of directing and identifying growth in knowledge of self and others, global attitudes, skills of interpreting and relating, skills of discovering and interacting, and critical cultural awareness. In a follow-up to his seminal work, Byram summarizes intercultural competencies as a more generalizable approach to global education with a focus on "knowledge of others; knowledge of self; skills to interpret and relate; skills to discover and to interact; valuing others' beliefs and behaviors; and relativizing oneself" (2021, p. 22). Researchers have noted that an increasingly globalized society has prioritized the need for cross-cultural skills development (Biraimah & Jotia, 2013; Parkhouse et al., 2016; Smolcic & Katunich, 2017). The desire for educators who demonstrate the capacity to respond to increased classroom diversity and global trends has shifted the development of intercultural competencies skills as a cross-curricular priority rather than a singular focus of language educators (Byram, 1997).

While Byram's (2021) framework provides a theoretical foundation for understanding intercultural competencies, it offers limited direction in building a construct for validating and assessing intercultural competencies, specifically concerning the skills of educators (Biautti, 2018; Deardorff, 2006; Morais & Ogden; 2011; Goren & Yemini 2017; Kerkhoff, 2017). Álvarez Valdivia and González Montoto (2018) challenge whether intercultural competencies can be measured as an individual attribute and suggest the distribution of these competencies throughout a learning community. Deardorff (2006) made an early effort to define intercultural competencies by utilizing the qualitative Delta method, allowing for structured anonymous communication to seek consensus among (n=23) university professors. The researcher found that preferred definitions were broad, creating challenges in developing an inventory for sufficient measurement.

Several themes have emerged from studies that have explored requisite potential attributes of intercultural competencies. Deardorff (2011) notes the importance of attitudes related to openness, respect, curiosity, and discovery as foundational steps leading to the development of skills, knowledge, and the potential for action. Fantini (2009, p. 459) identifies four similar dimensions related to attitude, knowledge, skills, and awareness. Educators who wish to foster global learning need to embrace differences, listen, and be open to new perspectives while withholding judgment (Baiutti, 2018; Deardorff, 2006; Portera, 2014). Curiosity differs from openness, emphasizing inquiry, cultural understanding, and appreciation for people with diverse backgrounds (Deardorff, 2006; Kerkhoff, 2017; Portera, 2014). The pursuit of discovery involves the willingness to take risks, embrace uncertainty, and explore diverse cultural perspectives (Deardorff, 2006; Kerkhoff, 2017; Kerkhoff & Cloud, 2020; Portera, 2014). These attitudes promote the capacity to develop the skills and knowledge that allow for an effective global learning environment to emerge.

Researchers emphasize the importance of well-developed knowledge, comprehension, and skills as essential components of intercultural competencies (Baiutti, 2018; Deardorff, 2006; Kerkhoff, 2017; Morais & Ogden, 2011). While language fluency has long been recognized as an essential skill for those wishing to develop a global perspective, it is insufficient on its own to make one interculturally competent (Deardorff, 2016; Fantini, 2009; Portera, 2014; Putman & Byker, 2020; Fantini, 2009, p. 459). Researchers identify adaptability and the ability to communicate effectively as the most universally essential indicators of intercultural competencies (Baiutti, 2018; Byram, 2021; Deardorff, 2006; Hammer, 2011; Kerkhoff, 2017; Morais & Ogden, 2011). A broad range of skills that impact intercultural connections includes adapting communication style, behaviors, and mannerisms while analyzing the context of the

situation and diverse perspectives represented (Deardorff, 2011; Fantini, 2009; Mansilla & Jackson, 2011; Morais & Ogden, 2011). Martorana et al. (2021) emphasize the importance of communicating constructively in diverse settings to reach mutually shared meaning and understanding. Modeling adaptability promotes related skill growth in students. Adaptability is an increasingly relevant skill for teachers as diversity in the classroom becomes more explicit and profound. Baiutti (2018, p. 562) notes that adaptable people demonstrate the ability to "decentralize themselves and recognize that their point of view is not universally valid." Adaptable educators possess the humility to suspend judgment based on their pre-established worldview and embrace the perspectives and experiences of others.

Attitudes and skills provide opportunities for knowledge of culture, traditions, and perspectives to be applied as a means of promoting a positive learning environment for globally diverse students (Biraimak & Jotia, 2013; Kerkhoff & Cloud, 2020; Martorana et al., 2021; Parkhouse et al.; 2016; Putman & Byker, 2020; Zhao, 2010). This knowledge enhances an appreciation for diverse learners while ensuring a cultural lens is applied when selecting curriculum materials (Biraimak & Jotia, 2013; Heizmann et al., 2015; van Werven et al., 2023). The insight gained through authentic knowledge is instrumental in helping to defuse stereotypes while validating diverse cultures and perspectives represented in the classroom (Alvarez Valdivia, 2018; Biraimak & Jotia, 2013; Deardorff, 2006; Kerkhoff & Cloud, 2020; Morais & Ogden, 2011; Zhao et al., 2009; Zhao, 2010). This knowledge, combined with the attitudes and skills, equip educators to lead initiatives that promote an inclusive and globally diverse environment while modeling for students a path for changes in the broader community.

Exploring the psychometric properties of intercultural competencies is limited due to the few existing studies examining quantitative instruments (Bustamante et al., 2017). The need for a

greater understanding of theory-based factors contributing to the overall construct of intercultural competencies elucidates a need to assess related skill levels. This is particularly true for teachers. The growth in professional development interventions, both domestically and abroad, requires a scalable approach that offers insight into longitudinal growth and implementation of new skills in the classroom. Developing a teacher-focused quantitative measurement instrument can help provide stronger validity and clarity to assessment data currently collected through qualitative means, including interviews, journals, reflection essays, and observation.

Existing Intercultural Competencies Assessment Instruments

The desire for a greater understanding of educators' intercultural competencies development has promoted research into developing a theoretically grounded scale and measurement instrument for assessing elements related to this construct. The Intercultural Development Inventory (IDI) is the most commonly applied intercultural competencies measurement instrument in schools, though it was not explicitly designed to measure the professional growth of teachers or have direct application in the classroom (He et al., 2017). The IDI consists of 50 multiple-choice items demonstrating high cross-cultural validity and reliability (Alvarez & Gonzalez Montoto, 2018; Hammer, 2011). The IDI assesses and positions participants on a 7-part developmental continuum-based ranging from the least developed stage, "denial," to the most advanced stage, "adaptation," with intermediate stages including "Polarization," "Defense," "Reversal," "Minimization," and Acceptance" (Alvarez & Gonzalez Montoto, 2018; Hammer, 2011; Jankowski, 2019). The IDI assesses participant responses and applies findings on measures in a continuum that scales a participant's ability to observe and take in cultural differences and commonalities while adapting one's behavior to the cultural context (Hammer, 2011; Jankowski, 2019). Heinzmann et al. (2015) argue that the chronology of the

developmental stages identified on the IDI continuum is not supported by adequate empirical evidence and that there is no certainty that passage through each stage is required. Morais and Ogden (2011) note that the IDI is limited in focusing on global competencies but is ineffective at measuring global citizenship. This potential conflict in context validity has motivated researchers to investigate more educationally aligned assessment instruments to promote intercultural competencies growth needs of educators.

Kerkhoff (2017) developed the Teaching for Global Readiness (TGR) scale for educators based on four identified dimensions: situated practice, integrated global learning, critical literacy instruction, and transactional experiences. The researcher utilized a sequential exploratory mixed methods approach involving qualitative interviews with (n=24) teachers to develop the TGR. This data was applied to exploratory factor analysis and confirmatory factor analysis to design a measurement scale with 19 items that satisfied these dimensions. The TGR differs from the IDI in focusing on the application of intercultural competencies to instructional practice rather than providing a scale that measures an individual's intercultural competencies (Hammer, 2011; Kerkhoff, 2017).

The UNESCO Program for International Student Assessment (PISA) also developed an assessment instrument related to pre-established international global competencies benchmarks (Chandir & Gorur, 2021; OEDC, 2020; Robertson, 2021). While designed for students, PISA offers insight into a framework for examining intercultural competencies in schools. PISA assessment was designed to be a triennial international survey for 15-year-old students, taking six-and-a-half hours to complete, and based on a scale targeting performance in mathematics, science, and reading (OECD, 2018). The 2018 PISA assessment included components that examined global competencies through a multi-dimensional construct that combined the domains

of *knowledge, skills, attitudes,* and *values* alongside four components: (a) examining local, global, and intercultural issues; (b) appreciating diverse global perspectives; (c) engaging in cross-cultural interaction and understanding diverse perspectives; (d) taking actions that support the collective well-being (OECD, 2020, p. 57). The 2018 PISA assessment consisted of two sections: (a) a *cognitive test* that assessed background knowledge and problem-solving skills related to global themes and (b) a *questionnaire* related to awareness of global issues, cultural understanding, attitudes, and information about how global competencies are taught in their schools (OECD, 2020, p. 17). An expert panel of university professors and thought leaders designed the framework and assessment items (Chandir & Gorur, 2021; Robertson, 2021). The cognitive assessment section consisted of 69 items organized into 18 units based on 4 clusters (OECD, 2020, p. 64). Items were based on a varied response format with a mix of open and closed responses (OECD, 2019, p. 194; OECD, 2020, p. 47). Chandir and Gorur (2021) noted clear western bias in the wording of items and found that sample items provided by PISA elicited varied responses depending on the cultural background of 15-year-old participants. While the United States was one of 50 countries that participated in the PISA assessment that opted out of the section related to global competencies, researchers note a U.S.-centered liberal political bias in assessment items, along with evidence of aiming to advance corporate interests" (Robertson, 2021). While PISA is a student assessment, it offered insight into a framework that could eventually inform how educator intercultural competencies are assessed.

While the IDI, TGR, and PISA scales provide theoretically driven approaches to measuring intercultural competencies that apply to educational settings, none of those scales specifically targets the intercultural competencies growth of educators in ways that benefit the classroom environment (Byram, 2021; Hammer, 2011; Kerkhoff, 2017; Zhao, 2010). The IDI

continuum-based rating provides insight into the individual's positioning, but this does not inform improvements in instructional practice. The TGR informs instructional practice but does not directly target personal growth. Byram (2021) warns of the limitations of using quantifiable assessments to measure intercultural competencies, noting that many competencies are too complex to summarize into scalable objectives. He argues that "complex competencies are assessed in complex ways," pointing out that critical concepts related to interaction and engagement are not measurable and that attempting to do so would be detrimental (Byram, 2021, p. 148). Current assessments of intercultural competence provide an incomplete picture with a focus on outcomes rather than the attitudes, behaviors, skills, and critical thought that define the process (Deardorff, 2016). Byram (2021) concedes that the advantages of using psychometric tests include the potential to serve the needs of the school environment and could outweigh the shortcomings. With these limitations in mind, researchers must provide insight into the measurement approach to intercultural competencies to further the continued professional growth of educators. Careful attention must be given to what specific areas of intercultural competencies should be tested and how these measures connect with instructional objectives (Fantini, 2009). The subsequent section will address these concerns, adding further insight into attributes of intercultural competencies that have already been defined. This section will also provide insight into how researchers respond to the reliability and validity of measurement instruments that aim to assess the intercultural competencies of educators.

Measurement Threats to Reliability and Validity

The construct validity of a scale ensures the intended interpretation of items by participants and the capacity and willingness of participants to provide accurate and objective responses (Huck, 2004). When assessing construct validity, researchers need to ensure the

measurement instrument is consistent with the desired theoretical construct (Hammer, 2011; Jankowski, 2019; Kerkhoff, 2017). Hammer (2011) and Kerkhoff (2017) both applied a confirmatory factor analysis (CFA) to check for model fit using results from exploratory samples. Kerkhoff (2017) used data collected from (*n*=289) participants to examine the factor structure of 30 proposed TGR items before examining CFA results and eliminating 11 items due to high standard residuals (six items) and low factor loading (five items). The researcher organized the remaining 19 items into four factors that provided a valid construct concerning measuring instructional practice. While a CFA is effective in determining model fit for a new measurement instrument, it can also be effective in better understanding more mature measures. Hammer (2011) examined the 50 items from the IDI Version 2, using data from (*n*=4763) participants and running a CFA to examine construct and content validity. Instead of using the CFA results to remove items, the researcher used the analysis results to include one additional scalable interval, cultural disengagement, that had not been previously identified on the measure.

The continued development of intercultural competencies assessment instruments and data collection from a more extensive and diverse sample of participants will promote a greater understanding of construct validity measures (Jankowski, 2019; Kerkhoff, 2017). Jankowski (2019) notes that analyzing construct validity is a never-ending process inevitably impacted by the diversity of participant groups and their related context. Awareness of this diversity promotes the purposeful development of items that better support the construct.

Impact of Context and Participant Diversity

Researchers have found that the importance of context is significant when determining the generalizability of intercultural competencies assessment instruments (Alvarez Valdivia & Gonzalez Montoto; 2018; Bernardes et al., 2019; Biraimah & Jotia, 2013; Choi & Shin, 2016;

Kerkhoff, 2017; Jankowski, 2019; Perry & Southwell, 2011). The external validity of an assessment instrument can be affected by a variety of factors, including the diversity of a school, curriculum standards, teacher autonomy, time devoted to professional development, and the intended use of the results (Choi & Shin, 2016; Okken et al., 2019). Following five years of qualitative observations of teachers engaged in pre-service field experience in Honduras, Malewski and Phillion (2009) noted the impact of class, gender, and race on shaping their perceptions of themselves and others. Okken et al. (2019) and Zhao et al. (2009) observed that school and home environments also impact how teachers' increased global understanding is applied in the classroom. Researchers have indicated that existing instruments, including the IDI and TGR, have only been used in limited settings and could benefit from more field-based tests in broader educational contexts (Hammer, 2011; Kerkhoff, 2017). Further development of intercultural competencies assessment instruments and a more comprehensive collection of data from participants will assist researchers in better understanding the impact of context while identifying the potential bias when making interpretations.

Participant Self-Reported Bias

Construct-irrelevant variance is an additional threat to validity that must be considered when analyzing results from self-reported assessment instruments (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). Most intercultural competencies assessment instruments rely on self-reported data impacted by inflated self-perceptions and conscious and subconscious bias (Alvarez Valdivia & Gonzalez Montoto, 2018; Kerkhoff, 2017; Morais and Ogden, 2011). Müller et al. (2020) identified that 88.6% of assessment tools for social, emotional, and intercultural competencies rely solely on self-reported data. Alvarez and Gonzalez Montoto

(2018) used the IDI as part of a mixed-method design to assess (*n*=8) elementary school teachers in Spain who were enrolled in an intercultural training activity. The researchers found participants' overestimation of their cultural sensitivity in the IDI data when compared to qualitative data collected during a focus group. When the potential for self-reporting bias is suspected, researchers must consider potential influences from external factors.

The perceived intended use of data collected from a measurement instrument can negatively affect validity. Kerkhoff (2017) notes that inflated self-perceptions on assessments can be minimized if teachers are confident that the data collected will not be used as part of an evaluation. The researcher also notes the need for alternative measures that can test the validity of self-reported data. However, the limited number of validated measures using a similar scale makes it challenging to identify and minimize the effect of construct-irrelevant variance.

Checking for concurrent validity by using multiple forms of assessments is one method researchers have identified to control for self-reporting bias (Kerkhoff, 2017; Jankowski,2019). Müller et al. (2020) recommend developing new intercultural competencies measurement instruments that rely on more than just self-reported data. He et al. (2017) noted the critical importance of using multiple measures when assessing (*n*=12) in-service teachers who participated in a 4-week professional development experience in Shanghai, China. Researchers compared IDI data to qualitative data from pre-and post-personal reflections and curriculum design projects created following the experience. The IDI results failed to measure faculty professional growth or application of global themes in the classroom, whereas the mixed methods approach identified heightened levels of empathy for students of diverse backgrounds and second language learners. The IDI is limited in that it is designed to place participants on a development scale rather than identify and assess teacher intercultural competencies that apply to

the classroom (Hammer, 2011). Therefore, current intercultural competencies assessment instruments provide the best control for self-reporting bias when incorporating the mixed methods approach. As more teacher-focused assessment instruments are developed, researchers will have more options for controlling concurrent validity and minimizing bias.

Longitudinal Assessments and Internal Validity

The development of a theory-driven assessment instrument for intercultural competencies will provide opportunities for teachers to understand better the longitudinal effects of interventions, including professional development opportunities, both abroad and domestically (Baiutti, 2018; Das et al., 2016; Driscoll et al., 2014; Jankowski,2019; Kambutu & Nganga, 2008; Kerkhoff, 2017; Morais & Ogden, 2011; Okken et al., 2019). Kerkhoff (2017) notes the value pre-and post-intervention assessments can provide towards further internal validity and evaluating the effectiveness of educational programs. Researchers have found that well-structured pre-and post-intervention reflection opportunities are critical components of global programs (Das et al., 2016; Driscoll et al., 2014; He et al., 2017; Parkhouse et al., 2016; Trilokekar & Kukar, 2011). Kambutu and Nganga (2008) recognized the importance of pre-and post-assessments in promoting reflections when examining pre-and post-program survey data from pre-and in-service educators who participated in immersion programs in Kenya. These researchers noted themes related to cultural transformation and greater critical self-reflection leading to an awareness of their privilege. An effective measurement instrument better ensures that intercultural competencies growth can be tracked over time and provide direction for future professional development and reflection.

Longitudinal data from an intercultural competencies assessment instrument can help uncover the meaningful presence and possible variations in sustained effect. This is of particular

importance that all teacher participants are exposed the same intervention, yet represent diversity in race, ethnicity, gender, global exposure, subject areas focus, grade level taught, and years of professional experience (Bernardes et al., 2019; Jankowski, 2019; Kambutu & Nganga, 2008; Trilokekar & Kukar, 2011). Researchers would better understand how these interventions impact pedagogical approaches, response to diversity in schools, and overall sense of self-efficacy (Das et al., 2016; Hur & Suh, 2018; Kerkhoff, 2017). In-service teachers' professional development experiences abroad are one increasingly popular intercultural competencies growth intervention method that could benefit from tracking the longitudinal effects measured through theory-driven teacher assessment instruments.

Increased in-service teacher professional development programs abroad provided a rationale for the continued effort to create intercultural competencies assessment instruments that will guide educators' personal and professional growth (Jankowski, 2019). Researchers note a need to understand better how professional development experiences abroad can enhance a teacher's worldview and lead to a greater emphasis on including global themes in the classroom in thoughtful and meaningful ways (Das et al., 2016; Biraimah & Jotia, 2013; O'Connor & Zeichner, 2011; Patterson, 2015). Researchers observe that educators participating in professional development experiences abroad are better equipped to think critically about pedagogical practices (Hur & Suh, 2018; Parkhouse et al., 2016). A longitudinal study, utilizing data from a purposeful measure, could assist in better understanding the relationship between the quality of design in a professional development experience abroad and the length of the program itself. (Das et al., 2016; Driscoll et al., 2014; Kambutu and Ngango, 2008; Malewski & Phillion, 2009; Okken et al., 2019). As schools consider the financial, timing, and logistical implications,

an increased understanding of the effect of professional development abroad programs will help guide their informed decisions.

Conclusion

Educators with intercultural competencies skills are in a position where they can empathetically address the needs of learners with diverse global backgrounds while promoting intercultural themes in the curriculum. They can develop effective solutions to educational challenges brought on by an increasingly globalized world. This literature review examined which intercultural competencies are essential to educators and how a theoretically grounded scale and measurement instrument could adequately assess educators' growth and inform instructional practice.

While the IDI measures individual intercultural competencies and the TGR scale targets instructional practice, there is yet no educator-focused instrument that combines these interlocking objectives. An assessment instrument that provides data for both individual intercultural competencies and applications to instructional practice would be highly desirable when targeting the professional growth of educators and evaluating the effects of related interventions. Parkhouse et al. (2016) emphasize that as teachers grow in intercultural competence, they become more engaged in global issues and committed to better understanding and incorporating 21st-century skills into the curriculum. Participation in professional development related to intercultural competencies growth, both domestically and abroad, can also influence an educator's personal development and foster independence and self-confidence (Biraimah & Jotia, 2013; Okken et al., 2019; Nielsen et al., 2019). As teachers navigate the complexity of teaching in culturally diverse environments while preparing students for entry into

an increasingly global society, attention to intercultural competencies growth can increase their self-efficacy and confidence in providing relevant and effective instruction.

The risk of self-reporting bias and the need to develop an assessment instrument adaptable to diverse contexts are challenging issues that must be addressed without increasing the timing or complexity of the measurement process. The relevance of an educator's intercultural competencies growth must also be communicated in ways that link to student growth and 21st-century skill development. Measurement research can help define constructs like "global citizenship" and "intercultural competencies" more clearly and utilize this understanding as a means of implementing more effective global education programs (Morais & Ogden, 2011). However, the results of any assessment are only as effective as the willingness of the teacher and institution to pursue professional development (Alvarez Valdivia & Gonzalez Montoto, 2018). The development of the ICME will aim to address instrument validity concerns while contributing to the broader understanding of how these professional competencies can be effectively measured with results applied in meaningful ways.

CHAPTER 3: METHOD

The purpose of this study was to develop the Intercultural Competency Measure for Educators (ICME) and examine its reliability and validity. This study examined the validity of a four-factor model based on the original theoretical latent constructs: (1) Intercultural Curriculum Content; (2) Intercultural Learning Environment; (3) Student Engagement; and (4) Personal Intercultural Competencies Growth. The data for this study was collected after the approval of the IRB in May and June 2022. This chapter will describe all the key aspects of the study, including the research questions, theoretical framework, instrument development, sampling participants, instrument design, data collection methods, and analytic approaches.

Research Questions

This study was structured by three interrelated research questions. The purpose of this scale was to measure intercultural competencies in K-12 educators. The first research question examined the content validity of the ICME by elucidating the logical link to the existing literature and feedback from experts in the respective field. The second research question examined the measure's construct validity through a confirmatory factor analysis. This tested the validity evidence of the four-factor theoretical model across the domains of intercultural competencies, attitudes, skills, knowledge, and actions. For the third research question, independent sample t-test and ANOVA were conducted to evaluate construct validity, examining if demographic characteristics known to be associated with intercultural competencies are related to scores obtained from the developed scale. These included: (a) proficiency in a second language; (b), subject area taught (c) experience living outside of one's primary citizenship, (d) race/ethnicity, and (e) gender. An independent Sample T-Test was used to examine continuous

independent variables while ANOVA was used to examine categorical and dichotomous independent variables. The following research questions guided the study:

1. *What is the extent to which content validity is established in the design of the Intercultural Competency Measure for Educators (ICME)?*
2. *To what extent does the ICME reveal construct validity evidence for its intended use?*
3. *To what extent do demographic characteristics commonly associated with contributing to intercultural competencies growth predict results of the ICME based on the four-factor theoretical model?*

Theoretical Framework

The theoretical framework for this study was based on critical cosmopolitanism (Byker, 2016; Kerkhoff, 2017; Müller et al., 2020; Oxley & Morris, 2013) as it relates to the understanding of intercultural competencies (Byram, 2021; Deardorff, 2006; Kerkhoff, 2017). Byram (1997) notes the importance of self-reflection as a foundational element in developing the *attitudes, skills, and knowledge* necessary for cultivating intercultural competencies. Deardorff (2006) acknowledges the work of Byram (1997) and highlights how these initial steps can eventually lead to *action* through a desire for social change. Critical cosmopolitanism calls for a heightened level of consciousness guided by a reflective process that leads to an enhanced awareness of political, moral, cultural, and economic relationships. This study utilizes commonalities between prior work on intercultural competencies and critical cosmopolitanism to develop four domains of intercultural competencies: (a) *Attitude*, (b) *Skills*, (c) *Knowledge*, and (d) *Action*. The first domain, Attitudes, reflects a person's respect, openness, curiosity, and tolerance for those who represent a different culture. The second domain, Skills, represents a person's ability to communicate, adapt, listen, empathize, observe, evaluate, or interpret when in

settings where cross-cultural differences exist. The third domain, Knowledge, is representative of a person's deep cultural insight, language fluency, and familiarity with current events. The fourth domain, Action, is representative of a person's willingness to advocate, change, improve, collaborate, or explore in ways that broaden cultural understanding or protect the interests of those who are marginalized due to cultural differences.

This study explored how the four domains of intercultural competencies cross with the four-factor theoretical model. These four theoretical factors were determined based on a review of the literature, results from the pilot study, and insight from the researcher's personal perception into educator intercultural competencies growth. The theoretical factor "*global curriculum content"* includes how classroom lessons and academic content reflect global and cultural themes. The theoretical factor "*intercultural learning environment"* reflects how the classroom and overall school setting reflect cultural inclusivity and global connections. The theoretical factor *"student engagement"* relates to how globally diverse students' learning needs, perspectives, voices, and contributions are recognized. The final theoretical factor "*personal intercultural competencies"* growth includes how participants explore, adapt to, and reflect upon new cultural experiences and engage with people of diverse cultural backgrounds.

Items were determined using the crossings for the four intercultural competencies domains and the theoretical four-factor model. This is illustrated in a matrix provided in Table 1.

Data Sources

Data collection for this study took place in May and June, 2022, during the final weeks of the 2021-2022 academic year. Data was collected using a convenience sample from participants at K-12 schools that are affiliated with the Global Education Benchmark Group

Table 1

Matrix Outline of the Four-Factor Intercultural Competencies Theoretical Model

Factor Area	Attitudes	Knowledge	Skills	Action
Intercultural Curriculum Content	Values the inclusion of global themes throughout the curriculum	Understands issues, topics, and diverse global perspectives	Able to adapt curriculum and evaluate student global learning growth	Promotes integration of cultural diversity through global curriculum
Intercultural Learning Environment	Values exploring perspectives and settings where all students feel represented	Understands strategies that build inclusive global learning environments	Able to ensure all feel valued by leveling cultural hierarchies	Promotes cultural diversity in the classroom and throughout the school
Student Engagement	Values the perspectives and contributions of globally diverse students	Understands how systematic inequalities impact student engagement	Able to adapt pedagogy to engage globally diverse students	Promotes the inclusion of all cultural backgrounds and perspectives
Personal Intercultural Competencies Growth	Values authentic exposure to unfamiliar cultures	Understands oneself through exposure to diverse cultures and people	Able to reflect internally and adapt when exposed to new cultural settings.	Promotes cross-cultural exposure, engagement, and understanding

(GEBG). The GEBG was established in 2007 as a non-profit organization consisting of nearly 308 member schools in 16 countries that, according to its mission, "researches and establishes model practices in the field of global education and supports member schools to prepare students to thrive in increasingly interconnected world systems" (GEBG Website, 2023, Vision Statement). GEBG member schools commit to emphasizing global education and citizenship throughout the school community, offering global education coursework, emphasizing second language proficiency, designating resources towards global education advancement, and purposefully engaging students in global opportunities. The researcher reached out to the

Executive Director of GEBG, who offered to provide a connection to member schools and encourage participation in the ICME study. Most GEBG schools are independent schools that have developed, or are developing, programs that promote K-12 global education. The researcher limited participation in this study to GEBG schools located in the United States and Canada. This allowed for greater continuity between participant schools.

Data Collection Procedures

Qualtrics was the primary tool for collecting data for this study. The 30 items on the ICME were ordered with introductory questions C.1, L.1, S.1, and P.1 appearing first. These attitude-based general questions allowed participants to adjust their thinking to the theme of the ICME before responding to items that were more specific. The remaining items were arranged so that the theoretical factor category and intercultural competencies domain were spread out. Having five or more response options generally helps increase the reliability of CFA results. All items were answered using a 6-point Likert Scale: (a) Strongly Disagree, (b) Disagree, (c) Slightly Disagree, (d) Slightly Agree, (e) Agree, and (f) Strongly Agree. The option of "neutral" was not offered to ensure participants were encouraged to provide a response for each item. The 6-point Likert scale allowed participants a range to more precisely express the degree to which they agreed or disagreed with each item.

Assessment Distribution Procedures

The researcher reached out to n=308 GEBG member schools in the United States and Canada with K-12 coeducational global education programs with a personalized email invitation to participate in the study. Schools accepting the invitation received a school-specific Qualtrics link for the ICME and introduction letter to be distributed to faculty. Most GEBG schools have a global education coordinator. This person, or another designated individual, coordinated the

distribution of the study in their respective school. It was recommended that the introductory letter and ICME be distributed during a faculty meeting and completed at that time. However, it was also possible for schools to distribute the ICME by electronic mail to be completed by a designated deadline. Prior to the deadline, the researcher updated the assessment distributor at each school with the number of forms completed along with a reminder of time remaining. Participants' data were collected in a password protected Qualtrics form file that was specific to each school. The school's name was coded on the forms and filed with a codebook kept on a separate computer that was password protected.

Once all data were collected, participant responses were compiled onto a common spreadsheet with all school identifiers removed. Individual school responses were processed into a collective item response summary for schools that meet the threshold of providing 30 responses or at least 50% of their faculty population. To protect participant privacy, the summary of item responses did not include information related to demographic data.

Participants

Participants in this study were K-12 educators at independent schools associated with GEBG in the United States. An objective of this study was to target in-service educators. These schools are often rigorous in their hiring process and attract highly qualified teachers who have demonstrated previous success in the classroom.

As noted in Research Question #3, this study examined how specific demographic characteristics are related to participant responses on the ICME. Participants were asked to respond to four questions regarding dichotomous demographic variables: (a) gender, (b) current teaching status, (c) second or foreign language proficiency, and (d) extended lived experience abroad for two or more years between ages 5 and 18 years. Participants were asked to respond to

three questions regarding categorical demographic variables: (a) core teaching level, (b) subject area taught, and (c) race/ethnicity. Participants were asked to respond to two questions regarding continuous demographic variables: (a) years in the education profession and (b) length of time lived abroad. Participants were also asked to identify if their school was public or independent, the general geographic region of their school, and if their school was rural, urban, or suburban.

Protection of Human Subjects

This study was submitted for IRB approval as it pertains to human subject research. For the purpose of being able to provide a summary overview to participating schools, the survey collected participant responses in school-specific data files, which included participant demographic information. To protect privacy, demographic data was withheld in summary reports provided to individual participating schools. Unique identifiers were kept for each school file, in place of actual school names, with the codes for these identifiers kept under password protection on a separate computer. Once all participant data were combined into a single data file, the unique identifiers for each school were removed. The researcher aggregated demographic data to protect the identities of participants that could be revealed using the information requested. While the personal data collected by this study is of low sensitivity, all data were kept in a secure location and password protected.

GEBG requested that each participating school that provided responses from at least 30 participants or at least 50% of its teaching faculty, whichever was lower, receive a general summary of the data collected. This allowed schools to analyze their own collective responses with the general responses provided by all schools. No individual participant responses were shared. Each participating school was provided with its own unique Qualtrics form of the ICME which appeared visually similar to all other forms. This made it easier to identify which school

was associated with each participant. For the purpose of this study, all participant information was gathered into one database with school identifiers removed.

Table 2

List and Description of Demographic Variables in the Model

Variable Name	Variable Type	Description
Participant Variables		
Teaching Students during the Current School Year	Dichotomous	0 = Yes; 1 = No
Years in Education	Continuous	0 = 4 years for less; 1 = 5-9 years; 2 = 10-14 years; 3 = 15-19 years; 4 = 20-24 years; 5 = 25 years or more
Core Teaching Level	Categorical	1 = Lower School Faculty; 2 = Middle School Faculty; 3 = Upper School Faculty; 4 = Non-teaching Staff; 5 = Non- Teaching Administrator; 6 = Other
Core Teaching Subject Area	Categorical	0 = Generalist (I teach all areas); 1 = Fine Arts (Arts, Music, Theater); 2 = Humanities (English, history, social studies); 3 = Math and Sciences; 4 = Modern & Classical Languages (Languages other than English); 5 = Physical Education; 6 = Special Support; 7 = Technology / Media Center 8 = Other; 9 = Non-Teaching Staff /Administrator
Proficiency in Second or Foreign Language	Dichotomous	0 = Yes; 1 = No
Length of Time Lived Outside of Country of Current Primary Citizenship	Continuous	0 = I have never lived in another country; 1 = Less than 3 months; 2 = 3-6 months; 3 = 7-11 months; 4 = 1-2 years; 5 = 3-5 years; 6 = 6-10 years; 7 = More than 10 years
Lived Outside of Country of Current Primary Citizenship for Two or More Years During Formative Years (between 5-18 years old)	Dichotomous	0 = Yes; 1 = No
Gender	Dichotomous	0 = male; 1 = female; 2 = other
Race/Ethnicity	Categorical	0 = American Indian or Alaska Native; 1 = Asian (including the Indian subcontinent) 2 = Black or of African Heritage; 3 = Hispanic/Latinx; 4 = Middle Eastern;

		5 = Native Hawaiian or Other Pacific Islander; 6 = White, non-Hispanic/Latino; 7 = Multiple Race/Ethnicities; 8 = Decline to Identify
School Variables		
School Type	Dichotomous	0 = Public; 0 = Private/Independent
Region of Current School	Categorical	0 = Northeastern United States; 1= Southeastern United States; 2= Southwestern United States; 3= Northwestern United States; 4 = Alaska/Hawaii; 5 = Canada; 6 = Latin America; 7 = Europe; 8 = Africa; 9 = Asia; 10 = Australia / Oceania
Community Environment	Categorical	0 = Urban; 1 = Suburban; 2 = Rural

Note.

Instrument Development

The literature provided insight into the validation and limitations of past assessment instruments designed to measure intercultural competencies in K-12 educators. The most commonly accepted intercultural competencies measure is the Intercultural Development Inventory (IDI) which offers a 50-item assessment and places participants on a seven-part developmental continuum (Alvarez Valdivia & Gonzalez Montoto, 2018; Hammer, 2011). While the IDI is used in educational settings, it was not explicitly created to address the needs of teachers or inform the work they do to facilitate student learning environments. There is a lack of evidence to support how results from the IDI contribute to the intercultural competencies development of K-12 educators or serve to enhance the learning needs of students. There is also a lack of adequate empirical evidence to support the chronology of the developmental stages identified on the IDI (Heinzmann et al., 2015).

Addressing shortcomings in the IDI, the Teaching for Global Readiness (TGR) scale was developed based on a cosmopolitan framework, with a focus on K-12 educators, and empirically validated with the following four dimensions identified: (a) situated practice, (b) integrated

global learning, (c) critical literacy instruction, and (d) transactional experiences (Kerkhoff, 2017). By addressing items related to skills and knowledge, the TGR offered insight into dimensions related to curriculum integration, learning environment, and student engagement that assist in informing factors theorized in this study. However, the TGR does not adequately address items related to behaviors and actions or examine factors related to educators' personal intercultural competencies growth. The TGR also targets classroom literature and technology and is less adaptable to all grade levels and subject areas related to a K-12 learning environment.

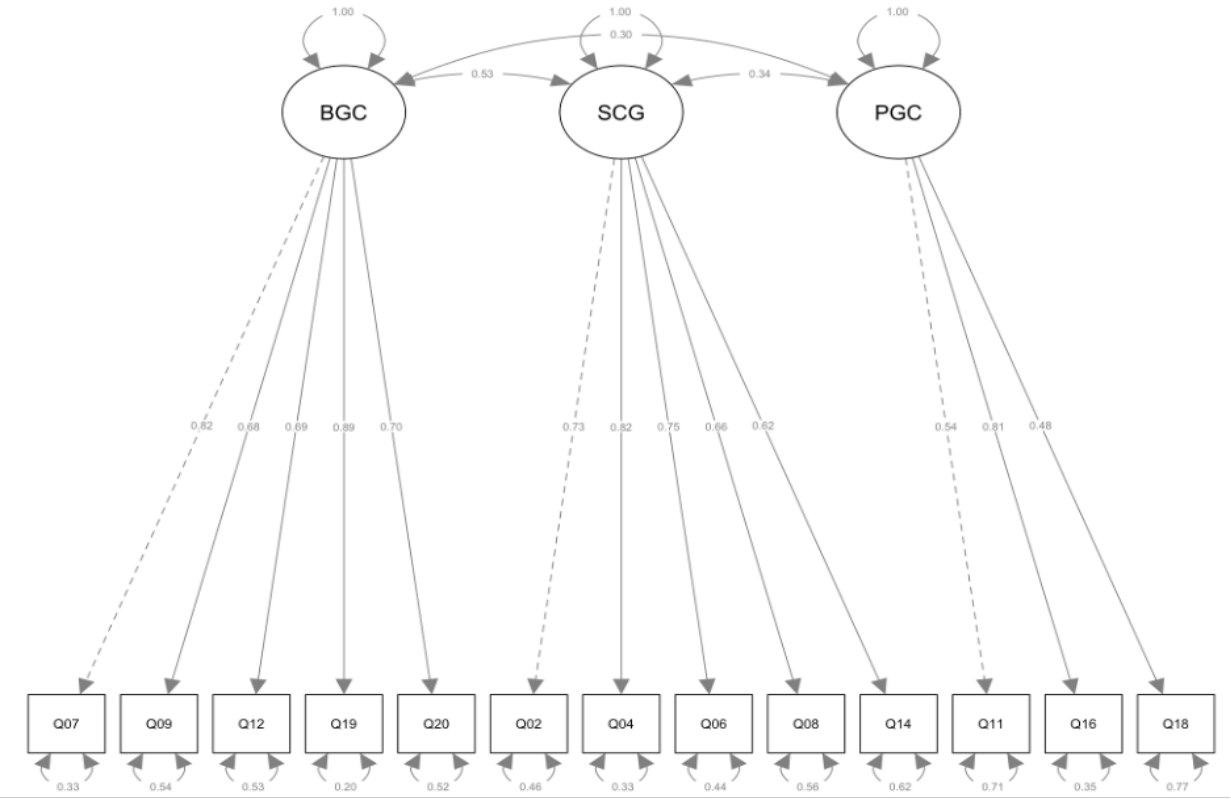
The creation of the ICME used in this study began with an extensive review of the literature. Based on the theoretical framework presented by Deardorff (2006), Kerkhoff (2017), and Baiutti (2018), a 20-item pilot assessment was completed by (n=102) educator participants at an independent K-12 school in a southeast metropolitan community. After outliers (n=5) were removed, participants demographics were 27% (n=25) male, 72% (n=66) female, 34% (n=31) elementary level teachers, 26% (n=24) middle school level teachers; and 33% (n=30) high school level teachers, and 8% (n=7) administrators who also taught. Zhou (2019) notes that validation analysis should be utilized throughout the scale development phase as an ongoing process of understanding how the scale is measured and construct defined. An exploratory factor analysis (EFA) was conducted only without a confirmatory factor analysis (CFA), recognizing that the total participant number was below the recommended minimum threshold of (n=100) (Streiner, 1994). An EFA showed three factors, with 13 out of the 20 items loading and accounting for 47.03% of the variability. A CFA of the three-factor model was tested based on the EFA and theory. Evidence was provided in the EFA and CFA that resulted in five items being removed. The CFA, based on the EFA of the pilot instrument, indicated that the initial model had a poor fit ($\chi 2$=401.744, df=170, p<.001, RMSEA=.117 [90% CI: .103 to .132];

CFI=.617; SRMR=.103). These results suggested multidimensional data be analyzed with a factor model. After recommended modifications were made, the final model showed an improved and more acceptable fit ($\chi 2$=75.215, df=62, p=.103; RMSEA=.049 [90% CI: .000 to .084]; CFI=.970; SRMR=.056). As shown in Figure 1, three distinct latent constructs were identified: (a) Building Global Curriculum, (b) Student Competencies Growth (respect, curiosity, global learning), and (c) Personal Global Competencies.

The three factors identified in the pilot study, including the 13 related items, were thoroughly analyzed along with a further review of the literature. The latent factors related to intercultural curriculum and personal intercultural competencies growth were theoretically supported with three or more items and maintained (Costello & Osborne, 2005). Yong and Pearce (2013) note that naming factors is an 'art' based on known variables within a factor. The latent construct of student competencies appeared to combine two distinct areas, one targeting student engagement in the learning environment and the other reflecting student engagement by educators. Items that did not load onto the pilot factor model were examined and it was observed that they could be rephrased to better fit the study's competencies factor. The researcher reflected on his professional experience in global education and discussed the factor with other professionals and the panel of experts. Considering this evidence and following a careful review of Kerkhoff (2017), the latent construct of student competencies growth was split into two factors related to the intercultural learning environment and student engagement. Items identified for each of the four theoretical factors were examined for fit with the four intercultural competencies domains: (a) Attitudes, (b) Skills, (c) Knowledge, and (d) Action. Additional items were considered to ensure that each theoretical factor included items related to each of the four intercultural competencies domains.

Figure 1

Factor Model from the Pilot Study of the ICME



Content Validation

Proposed items were assessed by the researcher for content validation related to clarity, brevity, and relevance. Grant and Davis (1997) and Cho and Cho (2017) note the importance of carefully selecting content experts during the scale review for content validation. Each of the 32 items was carefully examined by the researcher. The researcher assembled an expert panel with ten experienced global education practitioners. This expert panel included two university professors, four K-12 educators, three K-12 global education administrators, one head of school at a United States-based independent school, one head of school at an independent international school in eastern Asia, and one former head of school who now consults leaders at independent

international schools. Items were edited substantially for length and relevance. Multiple items were modified or eliminated due to redundancy. Thirty items related to the four theoretical factors and four intercultural competencies categories were included in the final version of the ICME (Table 2). The ICME was generally met with enthusiasm and support for the domains and factors. The expert panel provided insight into additional demographic factors that might be relevant when examining the relationship between the various predictors. While not an intended focus of the study, several expert panelists noted the scale's relevance to current school initiatives related to diversity, equity, and inclusion (DEI).

Analytic Approach

Factor analysis was used to inform the development of the ICME and to analyze the data collected. Gallagher and Brown (2013) suggested that it is appropriate to begin with an exploratory factor analysis (EFA), which uses data to create a model, prior to conducting a confirmatory factor analysis (CFA) which evaluates the theoretical model. An EFA was used to examine pilot instrument data and construct a theoretical model that served as a guide for enhancements made to the ICME. Once participant data was collected, a CFA was used to evaluate the hypothetical relationship between the manifest indicators (items) and latent variables (factors) (Gallagher & Brown, 2013). The examination of this relationship addressed the importance of analyzing the internal structure of the ICME to determine validity (AERA, APA, & NCME, 2014). The model was examined and the modification indices were reviewed for potential improvements.

Table 3

Items on the Intercultural Competency Measure for Educators (ICME)

Item		IC Domain
Intercultural Curriculum Content		
C.1	I value the inclusion of global content across the curriculum.	Attitudes
C.2	I believe that reaching second language proficiency is a priority.	Attitudes
C.3	I evaluate student global learning as it relates to my content area.	Skills
C.4	I adapt my curriculum content to include diverse global perspectives.	Skills
C.5	My knowledge of global issues and topics informs the content I teach.	Knowledge
C.6	My understanding of global perspectives is reflected in lessons I teach.	Knowledge
C.7	I discuss strategies for implementing global content with colleagues.	Action
C.8	I seek out and break down cultural stereotypes in curriculum.	Action
Intercultural Learning Environment		
L.1	I value settings where students explore diverse cultural perspectives.	Attitudes
L.2	I am curious how diverse student cultures can be better represented.	Attitudes
L.3	I create a classroom where culturally diverse students feel valued.	Skills
L.4	I identify cultural hierarchies that influence learning environments.	Skills
L.5	I know strategies for building inclusive global learning environments.	Knowledge
L.6	I ensure cultural diversity is clearly integrated into classroom settings.	Action
L.7	I collaborate with colleagues in promoting a culturally inclusive school.	Action
Student Engagement		
S.1	I am eager to learn about the cultures represented by my students.	Attitudes
S.2	I am open to new communication strategies that engage culturally diverse students.	Attitudes
S.3	I adapt pedagogy to engage culturally diverse students.	Skills
S.4	I am aware of ways cultural diversity impacts student engagement.	Knowledge
S.5	I recognize how systemic inequities inhibit the inclusion of culturally diverse learners.	Knowledge
S.6	I ensure culturally diverse students have a voice.	Action
S.7	I advocate for students whose home culture differs from societal norms.	Action
Personal Intercultural Competencies Growth		
P.1	I enjoy exploring unfamiliar cultures and people by reading books, watching films, or sampling unique ethnic foods.	Attitudes
P.2	When watching an international film, I prefer listening to words in the original language with subtitles rather than having words dubbed.	Attitudes
P.3	I adapt my behavior and mannerisms when engaging cross-culturally.	Skills
P.4	I first internally reflect, not judge, when cross-cultural challenges exist.	Skills
P.5	Through engaging with diverse cultures, I better understand my own.	Knowledge
P.6	I know how existing power structures impact culturally diverse people.	Knowledge
P.7	I actively support people marginalized due to cultural differences.	Action
P.8	I seek experiences that expand my cultural exposure.	Action

Note. All items will be answered using a 6-point Likert Scale: (a) Strongly Disagree, (b) Disagree, (c) Slightly Disagree, (d) Slightly Agree, (e) Agree, and (f) Strongly Agree

Since a sufficient sample size of the participants was collected, *n*=695, the data were randomly split using Excel software to allow the EFA and CFA to be conducted on separate samples. Using separate samples for the EFA and CFA helps to provide data that are more generalizable and avoid the risk of repeating the relationship in both analysis by using the same data (Bandalos & Finney, 2010, p.112). An EFA was conducted on the first half of the sample data followed by a CFA on the second half.

An Independent Sample T-Test was performed to examine the relationship between the ICME and demographic characteristics known to be associated with intercultural competencies. These characteristics included lived experience abroad, gender, race/ethnicity, subject area taught, and ability to speak a second language. A one-way ANOVA was used to examine this relationship when the demographic characteristic included three or more groups. These characteristics included length of time living outside one's primary country, grade level taught, and number of years teaching.

Exploratory Factor Analysis

An exploratory factor analysis was applied to examine the reliability and construct validity during the pilot stage of creating the ICME. Establishing reliability is critical when determining if results can be consistently reproduced (AERA, APA, & NCME, 2014). The software R was used to conduct the pilot analysis. An Oblimin rotation was used as it was determined there was a correlation between the subscales of the measure (Cho & Cho, 2017). Factors maintained were determined based on those that indicated an eigenvalue greater than one (Costello & Osborne, 2005). A Scree Test was also examined for an "elbow" to help determine the number of factors (Cattell, 1966). When reviewing the factors, items with a loading of less than 0.40 were deleted (Churchill, 1979; Gallagher & Brown, 2013). Cronbach's alpha was

reviewed to help determine reliability. A low Cronbach's alpha score indicates that the construct is not properly captured as being associated with the construct (Zhou, 2019). Correlation was reviewed to ensure at least 50% of the items were between 0.30 and 0.70 (Zhou, 2019). Correlations less than 0.30 were closely examined (Siwatu et al., 2017). The sample size of (n=98) was a shortcoming of the pilot analysis. Comrey and Lee (1992) recommend a sample size of 200 to 500 participants to run an EFA.

An analysis of the pilot study and EFA identified three factors. Further analysis and review of the literature theorized the existence of four factors and 30 items. Table 4 provides a structural model of the theoretical scale used in the ICME. A CFA was used to examine the construct validity of the internal structure of the ICME. Once the model specifications were determined by the EFA and data were collected, descriptive statistics were used to examine frequencies, means, and standard deviations of variables. The software R, using LAVAAN, were used to conduct the CFA (Rosseel, 2012). Gallagher and Brown (2013) recommend using maximum likelihood, determined by known model parameters, to account for missing data. Missing data were checked and identified for missing completely at random (MCAR), missing at random (MAR), or not missing at random (Gallagher and Brown, 2013). Mahalanobis distance was examined to check for multivariate outliers (Koran & Jaffari, 2020).

The adequate sample size was determined using the procedure provided by the Monte Carlo study. This was determined based on evidence of normality and conditions of missing data (Muthén & Muthén, 2002). The model also checked for correlation and linearity. A CFA formula examined the extent to which common variance is shared between the items making up the factor (Figure 2). The parameters of the model were examined for factor loadings, factor

variance/covariance, and error variance/covariance. This allowed for model identification and determined if there was sufficient evidence to estimate all free model parameters.

Figure 2

Confirmatory Factor Analysis Formula for the Four Factor Model

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \end{pmatrix} + \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} \eta + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{pmatrix}$$

Matrix: $\Sigma_y = \Lambda\Psi\Lambda' + \Theta_\epsilon$

Using the Shapiro–Wilk test, the p-value was examined for normality, helping to establish goodness of fit. With normality confirmed, Global Fit Indices were examined for model fit. Gallagher and Brown (2013) and Hu and Bentler (1999) note that the most commonly used indices for determining global fit include: (a) Chi-square test, (b) root mean square error of approximation (RMSEA), comparative fit index (CFI), (c) Tucker - Lewis index (TLI), and (d) the standardized root mean square residual (SRMR). However, the Chi-square test is often sensitive to sample size (Hu & Bentler, 1999). Goodness of fit helped determine the plausibility of the model.

Confirmatory Factor Analysis

The modification indices were examined for localized fit, highlighting potentially misidentified items and possible modifications. Modification recommendations greater than 3.84, greater than one degree of freedom, were examined and considered to see if they improved the model fit by freely estimating the parameter. However, modifications were checked to ensure they are theoretically and empirically grounded to avoid overfitting the model (Gallagher & Brown, 2013).

Independent Sample T-Test

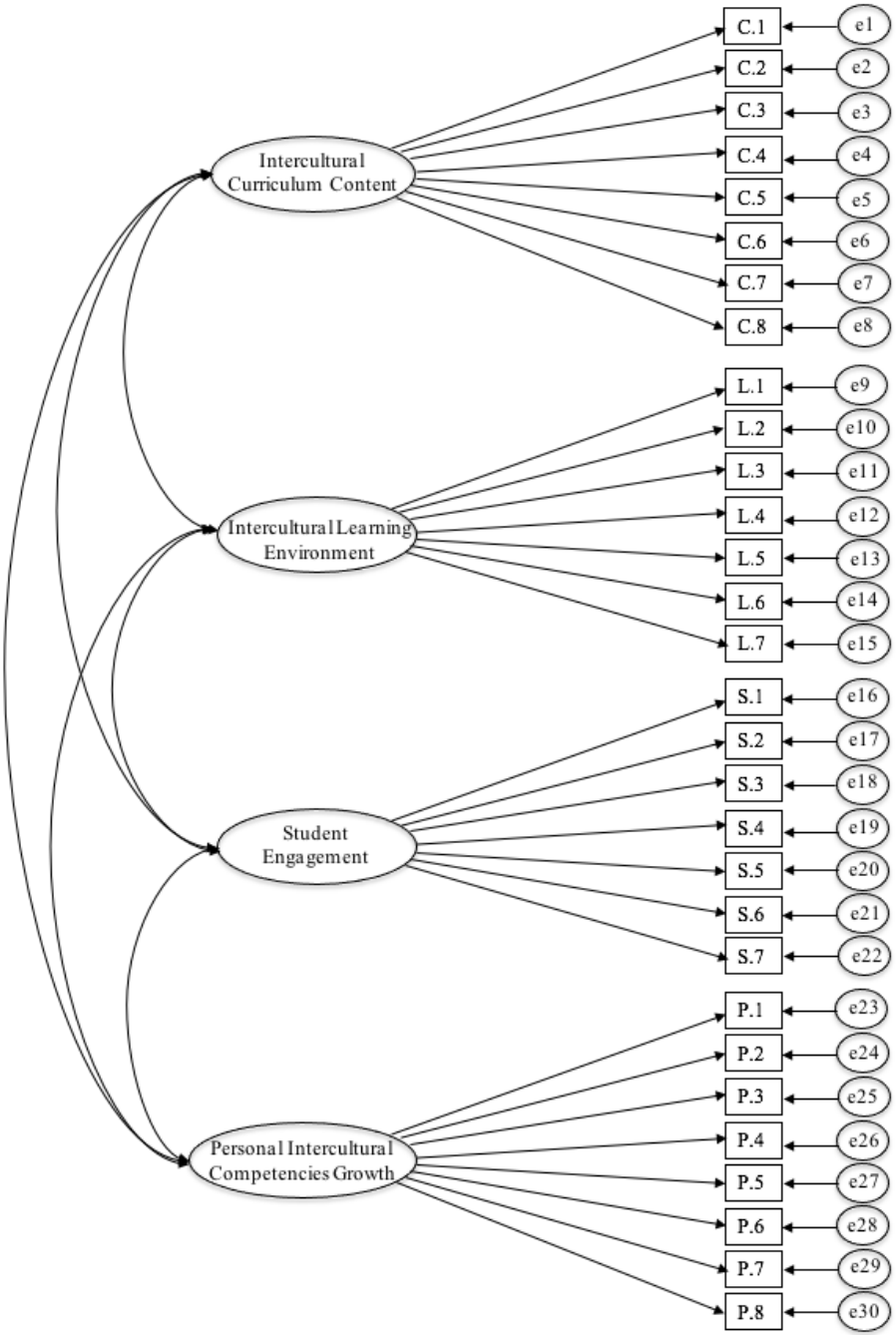
An Independent Sample T-Test examined how continuous demographic characteristics relate to the dependent variables found in the ICME. This was examined using SPSS Version 18 software. This data helped examine the relationship between the ICME and demographic characteristics, with two subgroups, known to be associated with educator intercultural competencies. Demographic categorical independent variables examined included the ability to speak a second language, race/ethnicity (white and people of color), subject area taught (Modern World Language and Math / Science), and gender. Hypotheses for each demographic characteristic were supported by evidence from the literature and discussions with the panel of experts. Effect size was also calculated using Cohen's D to help indicate how large the effect was when a statistically significant difference was found (Cohen, J, 1988, p. 12). The results of the Independent Sample T-test assisted in assessing for construct.

One-way ANOVA

A one-way Analysis of Variance (ANOVA) examined the extent to which categorical and dichotomous independent variables were compared based on the dependent variables found in the ICME. A one-way ANOVA helped examine the difference among the means to explore variations between and among subgroups. Demographic categorical independent variables examined included the length of time living outside one's primary country, grade level taught, and number of years as a teacher. Effect size was also calculated using partial eta-square to help indicate how large the effect was when a statistically significant difference was found (Tabachnick & Fidell, 2007, p.55). The results of the one-way ANOVA assisted in assessing for both construct and determinant validity.

Figure 3

Structural Model of the Intercultural Competency Measure for Educators (ICME) (Simple Structure)



Chapter IV: RESULTS

The purpose of this research study was to develop and validate an intercultural competencies scale for K-12 educators. This chapter will describe the results of data collected from (*n*=710) participants using exploratory factor analysis, confirmatory factor analysis, independent samples t-tests, and ANOVAs. This chapter will discuss the further refinement and validation of the ICME through the steps of data cleaning, reviewing descriptive statistics, analyzing statistical data, and examining validity. The review of findings and results addresses each of the research questions guiding this study:

1. *What is the extent to which content validity is established in the design of the Intercultural Competency Measure for Educators (ICME)?*
2. *To what extent does the ICME reveal construct validity evidence for its intended use?*
3. *To what extent do demographic characteristics commonly associated with contributing to intercultural competencies growth predict results of the ICME based on the four-factor theoretical model?*

Data Cleaning and File Preparation

Data were collected using Qualtrics Software over seven weeks between mid-May and the end of June 2022. Sixteen schools in the Global Education Benchmark Group (GEBG) network participated in the study, with data collected on identical Qualtrics forms created for each school. The (*n*=892) responses received by the end of June 2022 were exported to a common Microsoft Excel sheet. Participating schools were identified with a coded number in a school identity column alongside the collected data.

The initial Excel file of *n*=892 surveys collected was reviewed, with *n*=96 cases immediately eliminated due to completely blank responses. This was likely due to assessments

being opened and closed without completing any items. An additional *n*=53 (6.7%) cases were found to have more than 20% of their responses missing and were removed. There were *n*=33 cases where participants indicated that they were "non-teaching" administrators or staff. Since this study focuses on data collected from current K-12 educators, these participants were removed.

The remaining *n*=710 cases were reviewed for missing data by running a means imputation in SPSS 28. A Missing Values Analysis indicated that Little's test of Missing Completely at Random (MCAR) was significant, $\chi 2$=717.093, DF = 629, p = .008. This would reject the null hypothesis and indicate that data were not missing at random. While this is problematic, it is possible that the large sample size contributed to the significance. Since no item was missing more than *n*=6 responses (0.8%), the items missing for each participant was <20%, and the scales were constructed to allow for correlation between items, means substitution was used to impute missing data (Downey & King, 1998). Cases were randomly split into two groups using the "rand" function in Microsoft Excel. Each group consisted of *n*=355 cases and was stored in separate Excel files. An EFA was used to examine the first group to see if the factors identified matched the theoretical model. A CFA was used to examine the second group to see if factors identified from the first group provide a good model fit.

Description of Participant Schools

Sixteen schools, all part of the GEBG network, participated in the study. They responded to an electronic mail request to participate voluntarily. Schools with 30 or more participants were offered a 30-minute consultation session to review their schools' responses compared to the total population. Four other schools indicated a desire to participate but never followed through with teacher participation.

Table 4

Profile of Participating Schools

Participating School Demographics	*Number of Schools*	*Percentage of Schools*	*Faculty Participants*	*Percentage of Participant Responses*
Region				
Canada	2	12.5	137	19.3
Midwest - United States	2	12.5	41	5.8
Northeast - United States	6	37.5	278	39.2
Southeast - United States	3	18.8	206	29.0
West Coast - United States	2	12.5	48	6.8
School Size				
Less than 500 Students	3	18.8	91	12.8
501 to 1000 Students	8	50	376	53.0
1001 to 1500 Students	3	18.8	71	10.0
Over 1501 Students	2	12.5	172	24.2
Range in Grades Offers				
Jr. Kindergarten through 12	4	25	233	32.8
Kindergarten through 12	6	37.5	202	28.4
Grade 3 through 12	1	6.2	68	9.6
Grade 6 through 12	3	18.8	166	23.4
Grades 9 through 12	2	12.5	41	5.8
Public or Independent				
Independent	16	100	710	100.0
Urban, Suburban, Rural				
Suburban	16	100	710	100.0

Notes

Of the 16 participating schools (see Table 4), *n*=6 (37.5%) were in the Northeast; *n*=3 (18.8%) were in the Southeast; *n*=2 (12.5%) were on the West Coast; *n*=2 (12.5%) were in the Midwest; and *n*=2 (12.5%) were in Canada. According to Broughman et al. (2021), of the 30,492 independent schools in the United States, 22.1% are in the Northeast, 23.1% are in the Midwest, 35.0% are in the South, and 19.8% are in the West. Participating schools ranged in size, with the largest proportion *n*=8 (50%), having between 501 and 1000 students. All participating schools had secondary school programs with grades 9 through 12, and 10 (62.5%) schools educated students between kindergarten and grade 12. Like most schools in the GEBG network, all self-

reported as independent or private. All participating schools were located in suburban neighborhoods outside of major urban centers.

Description of Participants

Participants in this study were all in-service educators at GEBG member schools in the United States and Canada. Years in education ranged from *n*=145 (20.7 %), who were in the first ten years of their career, to *n*=165 (23.6), who had taught for 25 or more years. Most participants, *n*=367 (52.4%), taught at the upper school level (grades 9 through 12), which reflects that not all participating schools include younger grades. *n*=167 (23.9%) taught at the lower school level, *n*=132 (18.9%) taught at the middle school level, and *n*=34 (4.9%) reported as "other," likely teaching across multiple levels.

Participants representing core teaching subject areas that are often associated with intercultural competencies included *n*=171 (24.4%) humanities (English, history, and social studies); modern and classical languages, *n*=114 (16.3%); fine arts (arts, music, and drama), *n*=55 (7.9%); and generalists, n=98 (14.0%). Participants representing subject areas that are often less associated with intercultural competencies included math and science, *n*=148 (21.1%); physical education, *n*=22 (3.1%); special support, *n*=31 (4.4%); and technology/media center, n=18 (2.6%). Forty-three participants (6.1%) identified as teaching "other."

Of the participants who responded, *n*=291 (41.6%) reported being proficient in two or more languages, and *n*=85 (12.5%) indicated they had lived outside of their primary country of citizenship between the ages of 5 and 18. Most participants, *n*=390 (55.7%), had lived outside their primary country for at least three months, and *n*=166 (23.7%) had lived abroad for more than three years. This likely reflects the global education emphasis embraced at GEBG schools.

Table 5

Profile of Faculty Participants

Professional Experience and Demographics		*Number of Participants (Total)*	*Percentage of Participants (Total)*	*Number of Participants (Group A)*	*Percentage of Participants (Group A)*	*Number of Participants (Group B)*	*Percentage of Participants (Group B)*
Years in Education							
	4 years or less	44	6.3	25	7.2	19	5.5
	5 to 9 years	100	14.4	49	14.1	51	14.7
	10 to 14 years	112	16.1	54	15.6	58	16.7
	15 to 19 years	139	20.0	70	20.2	69	19.8
	20 to 24 years	135	19.4	59	17.0	76	22.0
	25 years or more	165	23.7	90	25.9	75	21.6
Core Teaching Level							
	Lower School Faculty	166	23.9	87	25.1	79	22.7
	Middle School Faculty	131	18.8	62	17.9	69	19.8
	Upper School Faculty	364	52.4	180	51.9	184	52.9
	Other	34	4.9	18	5.2	16	4.6
Core Teaching Subject Area							
	Generalist (teaching all areas)	98	14.1	52	15.0	46	13.2
	Fine Arts (Arts, Music, Theater)	55	7.9	23	6.6	32	9.2
	Humanities (English, History, Social Studies)	171	24.6	84	24.2	87	25.0
	Math and Sciences	146	21.0	76	21.9	70	20.1
	Modern & Classical Languages (Languages other than English)	112	16.1	55	15.9	57	16.4
	Physical Education	22	3.2	13	3.7	9	2.6
	Special Support	31	4.5	16	4.6	15	4.3
	Technology / Media Center	17	2.4	10	2.9	7	2.0
	Other	43	6.2	18	5.2	25	7.2
Proficient in Two or More Languages							
	Yes	288	41.4	138	39.8	150	43.1
	No	407	58.6	209	60.2	198	56.9
Lived outside of country of current primary citizenship for two or more years during formative years (between 5-18 years old)							
	Yes	85	12.2	45	13.0	40	11.5
	No	610	87.8	302	87.0	308	88.5
Gender							
	Male	196	28.3	102	29.4	94	27.0
	Female	481	69.2	238	68.6	243	69.8
	Non-binary / Third Gender	2	0.3	1	0.3	1	0.3
	Prefer not to identify	16	2.3	6	1.7	10	2.9

Professional Experience and Demographics	*Number of Participants (Total)*	*Percentage of Participants (Total)*	*Number of Participants (Group A)*	*Percentage of Participants (Group A)*	*Number of Participants (Group B)*	*Percentage of Participants (Group B)*
Race/Ethnicity						
American Indian or Native Alaskan	6	0.9	3	0.9	3	0.9
Asian (including the Indian Subcontinent)	27	3.9	15	4.3	12	3.4
Black or of African Heritage	38	5.5	27	7.8	11	3.2
Hispanic / Latinx	30	4.3	13	3.7	17	4.9
Middle Eastern	3	0.4	1	0.3	2	0.6
White, Non-Hispanic / Latino	535	77.0	264	76.1	271	77.9
Multiple Race / Ethnicities	20	2.9	10	2.9	10	2.9
Decline to Identify	36	5.2	14	4.0	22	6.3
Length of Time Lived Outside of Primary Country						
Never lived in another country	309	44.5	157	45.2	152	43.4
Less than 3 months	58	8.3	26	7.5	32	9.2
3 to 6 months	69	9.9	31	8.9	38	10.9
7 to 11 months	18	2.6	10	2.9	8	2.3
1 to 2 years	77	11.1	40	11.5	37	10.9
3 to 5 years	51	7.3	20	5.8	31	8.9
6 to 10 years	33	4.7	18	5.2	15	4.3
More than 10 years	80	11.5	45	13	35	10.1

Notes: Outliers (*n*=15) were removed (see EFA and CFA sections), results from *n*=695 participants were included in the study.

Reflective of the teaching profession, most participants reported being female, *n*=484 (69.1%), with *n*=18 (2.6%) identifying as non-binary/third gender or selecting not to identify. According to Taie and Goldring (2020), 74% of American independent school teachers identify as female. Seventy-seven (*n*=539) percent of the participants self-reported their race/ethnicity as white, Non-Hispanic/Latino. No other race/ethnic group was represented by more than 6% of the total participants. This included American Indian or Native American (1.0%), Asian (3.9%), Black/African Heritage (5.4%), Hispanic/Latinx (4.3%), Middle Eastern (0.4%), and multi-racial/ethnic (2.9%). 5.1% of participants chose not to identify their race or ethnicity. This is primarily reflective of overall national independent school teacher demographic in the United States: Native American (.4%), Asian (2.7%), Black/African Heritage (3.2%), Hispanic/ Latinx (7.2%), White, non-Hispanic (85.1%) and multi-racial/ethnic (1.3%) (Taie & Goldring, 2020).

Research Question One

The process used to address the first question, to what extent content validity is established in the design of the ICME, was examined throughout the research process. The inclusion of content experts, in addition to a careful review of the literature, is essential when establishing the content validity of a proposed scale (Grant & Davis, 1997; Cho & Cho, 2017). As noted in Chapter Three, the researcher assembled a research panel consisting of two university professors, four K-12 educators, three K-12 global education administrators, one head of school at a United States-based independent school, one head of school at an independent international school in eastern Asia, and one former head of school who now consults leaders at independent international schools. Carpenter (2017) notes the importance of having a diverse range of perspectives on a panel of experts, including representation of methodologists, intended participants, and subject matter researchers. Feedback from experts was utilized at each start of the scale development stage and during the validation process.

At the initial stage, once items were drafted based on the literate, expert feedback was used to assess the readability, quality, and over-related to the theme of the study. Feedback was provided through electronic mail, virtual discussions, and in-person consultation. After a pilot study was conducted with n=97 K12 faculty participants, the panel of experts screened results and research interpretations. One helpful recommendation was to include the domains of intercultural competence to categorize items alongside the factors where they loaded. Including a panel in the review of the makeup of factors aids in determining appropriate labels and reviewing connections that might otherwise be missed by the researcher (Carpenter, 2017). The panel aided in determining the ideal order and format for scale items, allowing for increased readability, avoidance of repetition, and the appropriate placement of challenging items (Fowler, 2014).

During the later stages of the study, the panel of experts was consulted to review the researcher's interpretation of CFA results. After two new factors emerged, "collaboration and adaptation" and "systematic awareness," the panel of experts helped to validate the statistical procedures and practical application of these areas. The items included in each factor were scrutinized as modifications were made during the CFA.

Research Question Two

The process used to address the second question, to what extent does the ICME reveal construct validity evidence for its intended use in assessing intercultural competencies in K-12 in-service educators, is described through the results and findings of both an exploratory factor analysis using Group 1 data and confirmatory factor analysis using Group 2 data.

Group 1: Descriptive Statistics and Correlations

Descriptive statistics were reviewed for each scale item based on the theoretical model. This examined each item's mean, standard deviation, skewness, and kurtosis (Table 6). The item responses were coded using a 6-point Likert Scale: 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Somewhat Disagree*), 4 (*Somewhat Agree*), 5 (*Agree*), and 6 (*Strongly Agree*).

Table 8

Descriptive Statistics for Group 1

Item	Response Category						Mean	SD	Skewness	Kurtosis
	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree				
L1	3	0	0	6	68	270	4.66	0.84	-3.97	18.34
S6	0	0	0	22	131	194	4.45	0.77	-2.41	10.25
C7	5	13	23	108	131	67	3.55	1.11	-0.89	0.96
P8	0	3	9	52	126	157	4.19	0.93	-1.44	2.98
L2	0	0	2	21	115	209	4.49	0.76	-2.30	8.98
P5	0	2	3	38	127	177	4.33	0.85	-1.81	5.29
C3	4	15	33	101	119	75	3.54	1.18	-0.77	0.39
L5	5	15	31	150	113	33	3.28	1.05	-0.76	1.10
C4	2	11	29	107	123	75	3.59	1.10	-0.74	0.58
P2	6	15	20	46	82	178	4.05	1.28	-1.41	1.29
S4	0	2	6	72	163	104	4.02	0.83	-0.86	1.61
P1	0	3	12	42	129	161	4.23	0.90	-1.35	2.17
C5	2	17	38	89	116	86	3.57	1.19	-0.69	0.03
L3	0	1	3	49	169	125	4.19	0.76	-0.99	2.45
S7	0	0	10	55	147	135	4.15	0.86	-1.04	1.60
C6	2	13	24	110	132	66	3.58	1.06	-0.78	0.77
P3	3	8	30	110	126	70	3.60	1.05	-0.69	0.60
L4	3	13	38	140	105	48	3.35	1.07	-0.50	0.43
P7	0	5	13	79	145	105	3.94	0.94	-0.96	1.42
C1	0	0	5	30	128	184	4.39	0.79	1.69	4.83
L6	1	6	22	87	156	75	3.75	0.98	-0.89	1.21
S5	1	0	6	59	145	136	4.14	0.88	-1.29	3.15
L7	2	8	15	109	128	85	3.73	1.03	-0.85	1.16
C8	1	9	31	89	130	87	3.69	1.10	-0.79	0.51
S1	0	0	1	22	101	223	4.54	0.71	-2.20	8.36
S2	0	0	0	19	126	202	4.50	0.69	-2.03	8.42
P6	0	3	12	76	144	112	3.97	0.93	0.97	1.46
C2	0	15	21	87	114	110	3.78	1.13	-0.83	0.29
S3	0	9	30	108	133	67	3.60	1.03	-0.61	0.31
P4	0	0	5	59	172	111	4.11	0.76	0.76	1.58

Note: Items are in the order in which they appeared on the ICME

The measures for skewness and kurtosis were reviewed for normal distribution, indicating if items are distributed in a way that will be meaningful to the final scale. Due to the nature of the scale and knowledge that participants were educators at GEBG-member schools, it was anticipated that there would be higher levels of kurtosis, a peak in the level of common

response, due to commonly shared global values. Eleven items indicated skewness and kurtosis of ≤ ± 1.0, potentially suggesting a normal distribution. Only one item had a skewness ≤ ± 2.5, item L1, “I value settings where students explore diverse cultural perspectives” (-3.97), reflecting the current educational prioritization of diversity, equity, and inclusion. Eight items had a kurtosis of ≤ ± 2.5, all indicating high peaks in levels of agreement. While measures for skewness and kurtosis challenge normal distribution assumptions, this was anticipated due to the sampling of educators participating in the study. All items with skewness and kurtosis of ≤ ± 2.5 were retained but noted for further consideration as factor analysis results were examined.

A Shapiro-Wilk normality test was performed and confirmed that the distribution of item responses departed significantly from normality ($W = 0.826$, p-value < 0.01). The rejection of the null hypothesis that all responses reflect normal distribution in the population promoted the use of the robust maximum likelihood estimation method during factor analysis.

A review of outliers in the Group 1 data set, using Mahalanobis distance, indicated eight outliers, cases 50, 130, 136, 165, 175, 222, 271, and 310. Upon examination of these outliers, all were found to have a streamline of consistent “highly disagree” responses. These cases were removed, leaving a total of n=347 cases.

A visual check of correlations in R indicated that most items correlated at ≥ 0.400. Items L1 and S6 indicated high levels of skewness and kurtosis, which were the only items that consistently correlated ≤ 0.400 with other items. The Kaiser-Meyer-Olkin (KMO) measure was also calculated to examine high and low correlations. The KMO = 0.95 indicated a high level of correlation between scale items, and >0.50 indicated a sufficient level of factorability (Yong & Pearce, 2013). Bartlett’s test of sphericity was also examined and determined to be statistically significant (p = >.001), indicating a significant difference in the variances to run factor analysis.

Exploratory Factor Analysis

A parallel analysis was first used to examine the recommended number of factors to verify if the four factors from the theoretical model were correct for the exploratory factor analysis (Table 7). A comparison of the reduced eigenvalues and 95 confidence intervals indicated the possibility of retaining four factors. A review of the Scree plot (Figure 4) revealed an "elbow" at four factors.

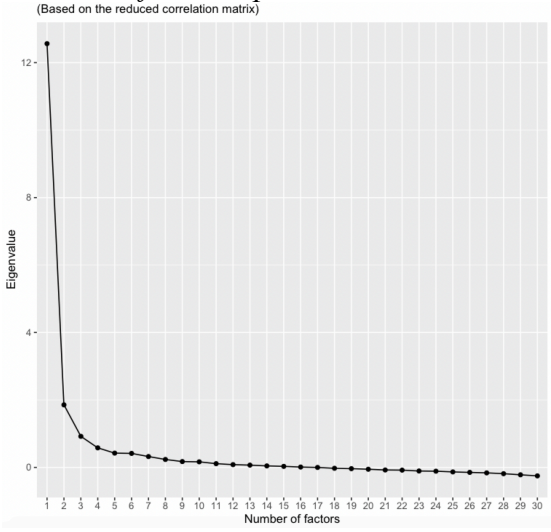
Table 7

Parallel Analysis for Group 1 Data

Factor	*Reduced Eigenvalue*	*Simulated Data Eigenvalue*	*95 Confidence Interval*
1	11.803	0.693	0.785
2	1.822	0.603	0.662
3	1.061	0.544	0.602
4	0.591	0.485	0.532
5	0.490	0.435	0.476
6	0.405	0.388	0.427
7	0.353	0.342	0.378

Figure 4

Scree Plot for Group 1 Data



An Exploratory Factor analysis (EFA) was run with four factors using an oblique rotation. A .40 loading cutoff was originally used with 26 items loading onto three factors. Yong and Pearce (2013) note that with a sample size >300, a factor of at least .32 would be needed to be considered statistically meaningful. After modifying the loading cutoff to .32 for the Group 1 data at four factors, 29 items loaded onto the three factors. An EFA with five factors was run with a .40 loading cutoff resulting in 26 items loading onto five factors. After modifying the loading cutoff to .32, all 29 items were able to load onto the five factors (Table 10).

Table 8

Factor Loadings from the Exploratory Factor Analysis

#	Item	*Factor 1*	*Factor 2*	*Factor 3*	*Factor 4*	*Factor 5*
C.5	My knowledge of global issues and topics informs the content I teach.	0.724				
C.6	My understanding of global perspectives is reflected in lessons I teach.	0.704				
P.1	I enjoy exploring unfamiliar cultures and people by reading books, watching films, or sampling unique ethnic foods.	0.589				
C.2	I believe that reaching second language proficiency is a priority.	0.556				
C.4	I adapt my curriculum content to include diverse global perspectives.	0.511				
P.2	When watching an international film, I prefer listening to words in the original language with subtitles rather than having words dubbed.	0.514				
C.3	I evaluate student global learning as it relates to my content area.	0.489				
P.3	I adapt my behavior and mannerisms when engaging cross-culturally.	0.445				
P.8	I seek experiences that expand my cultural exposure.	0.426				
L.4	I identify cultural hierarchies that influence learning environments.	0.420				
S.3	I adapt pedagogy to engage culturally diverse students.	0.333				
P.5	Through engaging with diverse cultures, I better understand my own.	0.323				
L.3	I create a classroom where culturally diverse students feel valued.		0.803			
S.6	I ensure culturally diverse students have a voice.		0.670			
S.7	I advocate for students whose home culture differs from societal norms.		0.545			
P.7	I actively support people marginalized due to cultural differences.		0.470			
L.6	I ensure cultural diversity is clearly integrated into classroom settings.		0.445			
P.4	I first internally reflect, not judge, when cross-cultural challenges exist.		0.387			
S.2	I am open to new communication strategies that engage culturally diverse students.			0.763		
S.1	I am eager to learn about the cultures represented by my students.			0.626		
L.2	I am curious how diverse student cultures can be better represented.			0.581		
C.1	I value the inclusion of global content across the curriculum.			0.425		
L.1	I value settings where students explore diverse cultural perspectives.					
C.7	I discuss strategies for implementing global content with colleagues.				0.761	
L.7	I collaborate with colleagues in promoting a culturally inclusive school.				0.623	
L.5	I know strategies for building inclusive global learning environments.				0.459	
C.8	I seek out and break down cultural stereotypes in curriculum.				0.406	0.333
S.4	I am aware of ways cultural diversity impacts student engagement.					0.666
P.6	I recognize how systemic inequities inhibit the inclusion of culturally diverse learners.					0.601
S.5	I know how existing power structures impact culturally diverse people.					0.418

Notes: EFA at 5 factors with a cut = 0.32

The one factor that failed to load, L1: *I value settings where students explore diverse and cultural perspectives*, also previously indicated high levels of skewness and kurtosis. This item was reviewed by the researcher and panel of experts and determined to be overly vague, and representative of a perspective commonly held by educators. This item was removed from further analysis.

One item, C8: *I seek out and break down cultural stereotypes in the curriculum*, loaded with two different factors. The researcher and panel of experts reviewed this item for further consideration after following a confirmatory factor analysis.

A four-factor EFA, consistent with the theoretical model, was run with items loading at .32. Three factors emerged from this EFA, with only three items, P6, S5, and C8, cross-loading into a fourth factor. These three items loaded poorly, >.390, and were determined to be a better fit for the other three factors. Two items, P4 and S4, failed to load in the four-factor EFA. The researcher and panel of experts reviewed the EFA results and agreed that the five-factor model was most favorable for continued analysis. The five-factor model included three factors that were similar to the theoretical four-factor model: (1) Curriculum and Understanding, (2) Student Engagement; and (3) Cross-Cultural Openness. The fourth item from the theoretical model, Learning Environments, was split into two new factors by the EFA results: (4) Collaboration Strategies and (5) Systematic Awareness. The new five-factor model was examined for percentage variance and accounted for 60.5% of the total variance (Table 9).

Table 9

Percentage Variance for Group 1 Data

Factor	*Eigenvalue*	*Percent Variance*	*Cumulative Variance*
1	12.24	40.81	40.81
2	2.27	7.57	48.37
3	1.57	5.23	53.60
4	1.09	3.63	57.23
5	0.99	3.29	60.52
6	0.94	3.13	63.65
7	0.91	3.03	66.68

The factor correlation matrix indicated a strong correlation between most of the factors, with weak correlations noted between factors 1 and 3 (0.273) and 3 and 4 (0.217) (Table 10). This justifies the use of the oblique rotation method. The structural correlation matrix strongly correlates the items and five factors, with a majority of the correlations measuring ≥0.400.

Table 10

Correlation between Factors in the EFA Five-Factor Model

Factor	*Factor 1*	*Factor 2*	*Factor 3*	*Factor 4*	*Factor 5*
1	1.000	0.490	0.273	0.633	0.446
2	0.490	1.000	0.364	0.544	0.363
3	0.273	0.364	1.000	0.217	0.494
4	0.633	0.544	0.217	1.000	0.375
5	0.446	0.363	0.494	0.375	1.000

Notes

Group 2: Descriptive Statistics and Correlations

Descriptive statistics were reviewed for each scale item based on the theoretical model. This examined each item's mean, standard deviation, skewness, and kurtosis (Table 11). The item responses were coded using a 6-point Likert Scale: 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Somewhat Disagree*), 4 (*Somewhat Agree*), 5 (*Agree*), and 6 (*Strongly Agree*).

Table 11

Descriptive Statistics for Group 2

Item	Response Category						Mean	SD	Skewness	Kurtosis
	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree				
L1	1	0	0	7	61	279	4.71	0.73	-4.20	22.59
S6	0	0	2	17	133	196	4.47	0.71	-2.01	8.11
C7	2	9	18	121	105	93	3.69	1.07	-0.70	0.64
P8	0	0	5	45	126	172	4.30	0.82	-1.37	3.20
L2	0	0	2	28	116	202	4.45	0.78	-1.91	5.95
P5	1	1	5	45	124	172	4.29	0.87	-1.60	4.14
C3	2	21	23	95	109	98	3.63	1.22	-0.84	0.29
L5	1	16	27	133	122	49	3.43	1.03	-0.61	0.56
C4	1	10	11	107	134	85	3.75	1.00	-0.79	0.96
P2	5	25	19	42	70	187	4.05	1.32	-1.34	0.78
S4	0	1	8	63	158	118	4.10	0.81	-0.70	0.42
P1	0	1	4	43	113	187	4.37	0.80	1.36	2.62
C5	2	12	20	108	112	94	3.71	1.08	-0.72	0.49
L3	0	0	2	39	158	149	4.29	0.72	-1.04	2.60
S7	0	2	1	59	147	139	4.19	0.82	-1.08	2.17
C6	0	4	28	95	128	93	3.78	0.98	-0.53	-0.04
P3	0	9	27	100	131	81	3.69	1.02	-0.60	0.14
L4	1	8	43	120	115	61	3.50	1.02	-0.32	-0.24
P7	0	6	4	71	163	104	4.00	0.87	-0.94	1.40
C1	0	1	3	29	112	203	4.45	0.76	-1.66	4.22
L6	0	6	21	83	143	95	3.85	0.97	-0.78	0.60
S5	0	2	6	45	155	140	4.21	0.81	1.17	2.44
L7	5	17	77	139	110		3.94	0.95	-0.82	0.70
C8	0	9	25	84	145	85	3.78	0.98	-0.66	0.14
S1	0	0	2	14	100	232	4.60	0.64	-2.11	7.70
S2	0	1	1	23	119	204	4.48	0.72	-1.74	5.13
P6	0	2	10	73	161	102	4.00	0.84	-0.79	1.20
C2	1	17	17	82	110	121	3.84	1.15	-0.94	0.44
S3	0	3	17	103	157	68	3.77	0.85	-0.38	-0.01
P4	0	0	5	49	190	104	4.11	0.74	-0.85	2.23

Note: Items are in the order in which they appeared on the Intercultural Competencies Scale

The measures for skewness and kurtosis for Group 2 were reviewed for normal distribution. Like Group 1, high levels of kurtosis were expected since participants taught at schools with a pre-existing commitment to global education. This was confirmed with ten items having a kurtosis of $\leq \pm 2.5$, all indicating high peaks in levels of agreement. Thirteen items

indicated skewness and kurtosis of ≤ ± 1.0. Only one item had a skewness ≤ ± 2.5, item L1, "I value settings where students explore diverse cultural perspectives" (-4.20), reflecting the current educational prioritization of diversity, equity, and inclusion. Due to the similar contexts in which participants in this study teach, the measures for skewness and kurtosis were anticipated, even while challenging normal distribution assumptions. All items with skewness and kurtosis of ≤ ± 2.5 were retained but noted as results from the confirmatory factor analysis were reviewed.

A Shapiro-Wilk normality test was performed and confirmed that the distribution of item responses departed significantly from normality ($W = 0.6883$, p-value < 0.01). The rejection of the null hypothesis that all responses reflect normal distribution in the population promoted the use of the robust maximum likelihood estimation method during factor analysis.

A review of outliers in the Group 2 data set, using Mahalanobis distance, indicated seven outliers, cases 45, 61, 117, 138, 255, 236, and 270. Upon examination of these outliers, all were found to have a streamline of consistent "highly disagree" responses. These cases were removed, leaving a total of n=348 cases in Group 2.

A visual check of correlations in R indicated that most items correlated at ≥0.400. Items L1 and S6 indicated high levels of skewness and kurtosis, which were the only items that consistently correlated ≤0.400 with other items. The Kaiser-Meyer-Olkin (KMO) measure was also calculated to examine high and low correlations. The KMO = 0.95 indicated a high level of correlation between scale items, and >0.60 indicated a sufficient level of factorability. Bartlett's test of sphericity was also examined and determined to be statistically significant (p = >.001), indicating a significant difference in the variances.

Confirmatory Factor Analysis

A confirmatory factor analysis of the five-factor model identified by the EFA using Group 1 data was conducted using data from Group 2. A single-factor model was first examined using the Global Fit Indices for model fit (Figure 5). Gallagher and Brown (2013) noted that the most commonly used indices include: (a) Chi-square test, (b) root mean square error of approximation (RMSEA), comparative fit index (CFI), (c) Tucker-Lewis index (TLI), and (d) the standardized root mean square residual (SRMR). The Chi-square test is often sensitive to sample size, making it less reliable for this study (Hu & Bentler, 1999). An SRMR, typically between 0-1, of less than .08, indicates an acceptable fit (Gallagher & Brown, 2013). The RMSEA, which can be scaled from zero to infinity, but rarely exceeds 1, should be less than .06 to indicate an acceptable fit (Hu & Bentler, 1999). The CFI, also scaled 0 – 1, should be greater than 0.9 to indicate proper fit (Hu & Bentler, 1999). The TLI should be close to 1, or at least greater than 0.9 or .95, to indicate a good fit. (Hu & Bentler, 1999). For the Group 2 single-factor model, the model fit was poor (χ2=1393.786, df=405, p<.001, RMSEA=.084 [90% CI: .079 to .088]; CFI=0.780; TLI=0.763; SRMR=.072). This justified further exploration of the five-factor model.

The original theoretical four-factor model was also examined using a CFA with data from Group 2. The result was a better model fit than the single factor model (χ2=1154.450, df=399, p<.001, RMSEA=.079 [90% CI: .074 to .085]; CFI=0.833; TLI=0.817; SRMR=.072). The five-factor model determined based on Group 1 data in the EFA was also examined using a CFA with data from Group 2. The result was a better model fit than the single-factor model and a slightly better model fit than the original theoretical four-factor model (χ2=1052.084, df=367, p<.001, RMSEA=.076 [90% CI: .068 to .078]; CFI=0.844; TLI=0.827; SRMR=.072) (Table 14). The initial results of the EFA and CFA, as well as the original four-factor theoretical model and new

five-factor model, were reviewed by the researcher and Panel of Experts. The decision was made to further explore the five-factor based on emerging factors related to collaboration and systematic awareness.

Figure 5

Single Factor Model for Group 2 Data

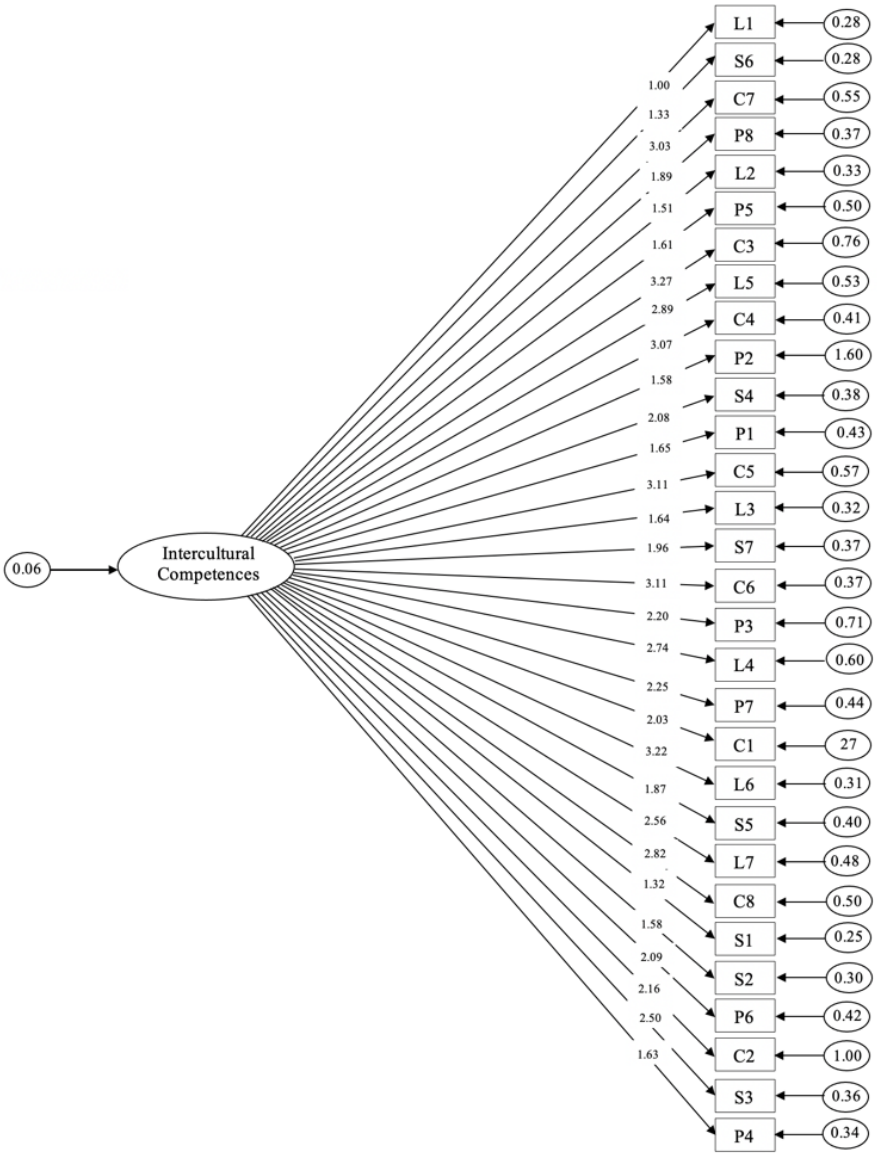


Table 12

Comparing the Model Fit using Group 2 data of the Single Factor Model, Four-Factor Theoretical Model and New Five-Factor Model

	$\chi 2$	*RMSEA*	*SRMR*	*CFI*	*TLI*
Single Factor Model	1393.786, df=405, p<.001	0.084	0.072	0.780	0.763
Four-Factor Theoretical Model	1154.450, df=399, p<.001	0.079	0.072	0.833	0.817
Five-Factor Model based on Group 1 data in the EFA	1052.084, df=367, p<.001	0.076	0.072	0.844	0.827

Notes:

Examination of the five-factor model alongside recommendations from the modification indices (MI) was reviewed for adjustments to the model that might improve the fit and further distinguish each of the five factors. When reviewing the MI recommendations, only suggestions with indices >10.00 were considered due to their likelihood of affecting overall model fit (Byrne, B. M., 2016). Modifications were carefully considered for effect on the overall theme of each factor and the panel of experts reviewed the results.

Ten modifications were applied to clarify and strengthen the themes for each of the five factors while also improving the overall model fit ($\chi 2$=1098.096, df=395, p<.001, RMSEA=.071 [90% CI: .067 to .076]; CFI=0.845; TLI=0.829; SRMR=.069) (Table 13). Four items, L1, P4, P5, and P8, were eliminated after they exhibited a weak factor loading in the EFA and were poorly linked to the five new factors. Item L1, *I value settings where students explore diverse cultural perspectives,* failed to load (<3.20) in the EFA, and exhibited high levels of skewness (-4.20) and kurtosis (22.59), violating normal distribution. Item P4, *I first internally reflect, not judge, when cross-cultural challenges exist*, had a low EFA loading (.387), and required a self-assessment of each participant's decision-making process that was perhaps challenging to gauge

in a cross-cultural environment. Item P5, *through engagement with diverse cultures, I better understand my own,* also had a low EFA loading (.323) and was perhaps too complex in asking participants to reflect on how they shape their understanding of their own culture. Item P8, *I seek experiences that expand on my cultural exposure,* had the highest loading (.426) of items removed but was perhaps vague in specifying what types of experiences were referenced. Three of the four items eliminated were from the *Personal Intercultural Competencies Growth* factor in the original four-factor theoretical model. These three items encourage participants to reflect on personal behaviors and do not explicitly relate to the school environment or student interaction. The decision to remove these items was reviewed with the panel of experts that found them to not fit well with the themes of the new five-factor model. Eliminating these items improved the overall model fit and supported the decision to remove them from the model.

A fifth item, P2, *When watching an international film, I prefer listening to words in the original language with subtitles rather than having words dubbed*, was removed when it was found to correlate with item P1, *I enjoy exploring unfamiliar cultures and people by reading books, watching films, or sampling unique ethnic foods*. P1 had a higher loading, .589 compared to .514, and was found to offer greater clarity and connection with the theme of the scale when reviewed by a panel of experts. Two additional sets of items were found to correlate, S1 and S2, and L3 and S7. The decision to retain both pairs of correlated items was based on the uniqueness of each item, how each correlated pair loaded under the same factor, and the items in each pair contributed distinctly to the factor theme. One item, C8, *I seek to break down cultural stereotypes in curriculum*, cross-loaded in the EFA into two factors. This item was placed under the factor *Collaboration and Adaptation*, where it had initially loaded slightly higher (.406) and better contributed to the overall factor theme.

Table 13

CFA Modifications Made to the Intercultural Competencies Assessment Scale

	$\chi 2$	*RMSEA*	*SRMR*	*CFI*	*TLI*	*Modification*
Based on EFA	1052.084, df=367, p<.001	0.076	0.072	0.844	0.827	
Round 1	869.331, df=289, p<.001	0.076	0.073	0.854	0.836	Removed items P5, P4, P8, and L1
Round 2	819.045, df=289, p<.001	0.073	0.072	0.864	0.851	Moved item L4 to Factor 5
Round 3	767.287, df=289, p<.001	0.069	0.069	0.880	0.865	Moved item S3 to Factor 4
Round 4	729.924, df=289, p<.001	0.066	0.064	0.888	0.875	Moved item P1 to Factor 3
Round 5	661.734, df=265, p<.001	0.066	0.062	0.898	0.884	Correlated items P1 and P2; Removed P2
Round 6	641.944, df=265, p<.001	0.064	0.061	0.902	0.890	Moved item P7 to Factor 5
Round 7	615.376, df=264, p<.001	0.062	0.059	0.909	0.897	Correlated items S1 and S2
Round 8	591.705, df=263, p<.001	0.060	0.059	0.913	0.903	Correlated items L3 and S7
Round 9	581.121, df=263, p<.001	0.059	0.057	0.918	0.906	Moved item P3 to Factor 4
Round 10	586.442, df=263, p<.001	0.059	0.057	0.916	0.904	Moved item L5 to Factor 2

CFA Modification recommendations were considered and made to move two items to the newly created factor *Systematic Awareness*. The items, L4, *I identify cultural hierarchies that*

influence learning environments, and P7, *I actively support people marginalized due to cultural differences*, both offered insight into how systems influence learning and limit the capacity of individuals due to their cultural backgrounds. Two recommended modifications were considered and made to move two items to the newly created factor *Collaboration and Adaptation*. The items, P3, *I adapt my behavior and mannerisms when engaging cross-culturally*, and S3, *I adapt pedagogy to engage culturally diverse students*, both address the needs of educators to be flexible and amendable to diverse settings and needs. Item P1, *I enjoy exploring unfamiliar cultures and people by reading books, watching films, or sampling unique ethnic foods*, from *Curriculum* was moved to *Cross-Cultural Openness* where it better aligned with the theme of the factor. A final modification recommendation was considered that transferred item L5, *I know strategies for building inclusive learning environments*, from *Collaboration and Adaptation* to "Diverse Student Inclusion" due to the important role of inclusive learning environments in engaging diverse students. These modifications helped to improve the overall model fit and were reviewed by the panel of experts to ensure they supported the theme of the individual factors and overall scale. The final five-factor model provided a balance of five items per factor (Figure 6). The five factors were identified as: (1) Curriculum, (2) Student Inclusion, (3) Cross-Cultural Openness, (4) Collaboration and Adaptation, and (5) Systematic Awareness.

Discriminant and Convergent Validity

The correlations between factors were compared for discriminant validity and convergent validity (Gallagher & Brown, 2013). It was expected that the correlation between factors would be high because that ICME was based on the theory. In examining the five factors, high correlations were noted between multiple factors (Table 14). The factor *Collaboration and Adaptation* had the highest correlation with other factors including *Curriculum* (0.904),

Systematic Awareness (0.884), and *Diverse Student Inclusion* (1.004). This raised concerns that discriminant validity might be violated and that this factor may not be measuring distinctly different characteristics from other factors (Gallagher & Brown, 2013). This could also be attributed to the limited number of items used to measure each factor. The only minor challenge to convergent validity was between the factors of curriculum and diverse student engagement (0.490). Overall, five factor correlations demonstrated convergent validity and likely measured the same latent variable of educators' intercultural competencies.

Table 14

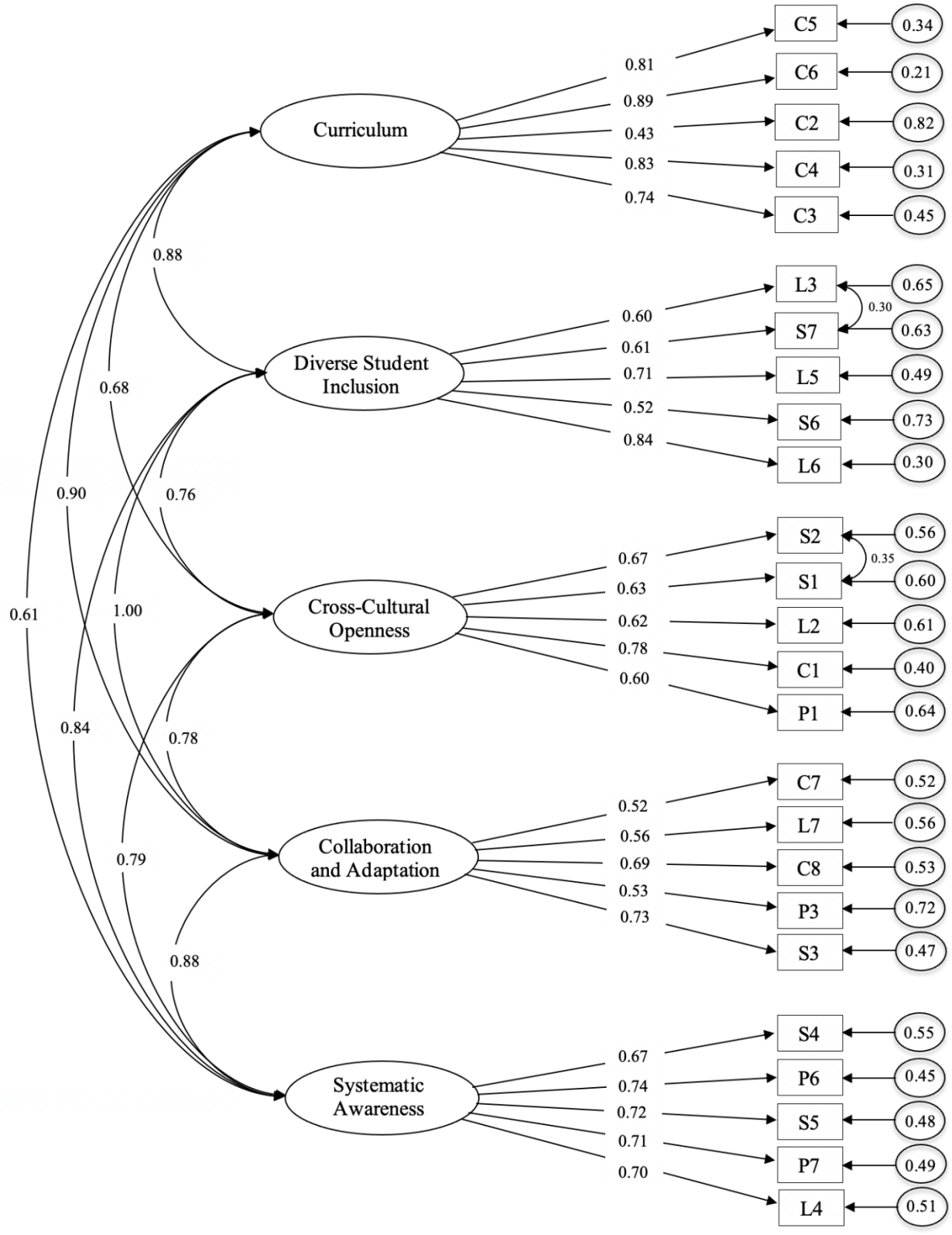
Correlation Between Factors in the CFA Five-Factor Model

Factor	*Curriculum*	*Diverse Student Inclusion*	*Cross-Cultural Openness*	*Collaboration and Adaptation*	*Systematic Awareness*
Curriculum	1.000	0.884	0.684	0.904	0.610
Diverse Student Inclusion	0.490	1.000	0.762	1.004	0.835
Cross-Cultural Openness	0.684	0.762	1.000	0.784	0.792
Collaboration and Adaptation	0.904	1.004	0.884	1.000	0.884
Systematic Awareness	0.610	0.835	0.792	0.884	1.000

Notes

Figure 6

Structural Model of the Intercultural Competency Measure for Educators (ICME)



Research Question Three

The final question of the study examines the extent to which demographic characteristics commonly associated with contributing to intercultural competencies growth predict the results of the ICME based on the five-factor model. Results showing significant correlations between ratings from the ICME and external variables that are expected to be associated will further strengthen the validity argument for the ICME. An independent sample t-test, conducted in SPSS 28, was used when comparing two distinct demographic categorical independent variables, including the ability to speak a second language, race/ethnicity (white and people of color), subject area taught (Modern World Language and Math /Science), and gender. The sampling was checked for outliers and normal distribution, while noting the likelihood that the homogenous sample population increased the risk of skewness and kurtosis. Homogeneity of variance was checked using Levene's Test of Equality of Variance with a significance level set at $p = 0.05$ (Gastwirth et al., 2009). When multiple comparisons were used, Bonferroni correction was applied, adjusting the p-value to decrease the likelihood of type I errors (VanderWeele & Mathur, 2019). When homogeneity of variance was not satisfied, Welch's t-test was used in the interpretation of results (Delacre et al., 2017). Whenever possible, groups with roughly equal sample size were examined. The dependent variables, items from the ICME, use Likert scale with six response options, allowing the results to be treated as scaled data (Dawes, 2008). The effect size was calculated following each independent t-test, using Cohen's D, to determine if significance found in the t-test relates to a small, medium, or large effect between the two group means and the dependent variable (Cohen, J, 1988, p. 12).

An ANOVA, conducted in SPSS 28, was used to compare means between three or more independent groups. Categories selected included subject areas taught, length of time in

education, and grade level taught. Assumptions were checked using similar procedures from the independent t-test for the dependent and independent variables. This included checking for roughly equal sample sizes, outliers, and normal distribution. Levene's test of equality of variance was used to verify homogeneity of variance. When homogeneity of variance was not satisfied, Welch's ANOVA was examined. An effect size, determined using partial eta squared, was calculated when an ANOVA test was significant to examine the strength of association between the independent and dependent variables (Tabachnick & Fidell, 2007). These outcomes from the independent t-tests and ANOVAs were used to determine if the ICME was assessing what it was designed to measure. The results of the both tests assisted in examining both construct and determinant validity.

Ability to Speak a Second Language

An independent-samples t-test was conducted to compare responses from teachers who are proficient in a second language and those who speak only English. Second language proficiency is often recognized as a critical disposition for educator global competence (Tichnor-Wagner et al., 2016; Zhao, 2010). Proficiency in a second language is recognized as an important 21st Century Skill and a key element when building intercultural relationships (Byker, 2016). The researcher hypothesized that responses on the ICME would indicate a moderate statistically significant difference with participants who spoke a second language having a higher mean average than those who spoke only English. Prior to running the independent t-test, an examination of Levene's test demonstrated that the assumption of homogeneity of variance was met, $p = .409$. Results indicated that the groups differed significantly on the overall assessment $t(653) = 7.427$, $p < .001$, $d = .572$. The means for faculty who spoke two or more languages (M=4.164, SD=0.556) was significantly different from faculty who were monolingual (M=3.840,

SD=0.573). Faculty who spoke two or more languages had a significantly higher mean average than those who spoke only English for all five factors (Table 15). This difference was most apparent for "curriculum", $t(653) = 10.39$, $p < .001$, $d = .800$. The means for faculty who speak two or more languages (M=4.079, SD=0.764) was significantly different from faculty who speak only English (M=3.436, SD=0.829). This is likely due to the importance of knowledge of multiple languages in subject areas like modern world language and the humanities. These findings align with the hypothesis that those who speak multiple languages have a higher level of intercultural competence and supports the construct and determinant validity of the ICME.

Table 15

Independent Sample t-test for Ability to Speak a Second Language

	Yes (*n*=288)	No (*n*=407)						
Factor	*M/SD*	*M/SD*	*F*	*Sig.*	*t*	*df*	*p*	*d*
Curriculum	M=4.079 SD=0.764	M=3.436 SD=0.829	2.192	0.139	10.39	693	<.001	0.800
Student Inclusion	M=4.172 SD=0.601	M=3.924 SD=0.624	1.708	0.192	5.238	693	<.001	0.403
Cross-Cultural Openness	M=4.580 SD=0.503	M=4.403 SD=0.507	0.381	0.537	4.559	693	<.001	0.351
Collaboration and Adaptation	M=3.930 SD=0.713	M=3.578 SD=0.737	0.261	0.610	6.283	693	<.001	0.484
Systematic Awareness	M=4.058 SD=0.647	M=3.858 SD=0.661	0.686	0.408	3.957	693	<.001	0.305
Total Scale	M=4.164 SD=0.556	M=3.840 SD=0.573	0.683	0.409	7.427	693	<.001	0.572

Notes: Bonferroni correction was applied, setting the p-value for significance at <.008.

Modern World Language and Math/Science Teachers

An independent-samples t-test was conducted to compare responses from participants who teach modern world languages with participants who teach math and science. The MWL and MS were selected as demographic categories to compare from a broader range of curriculum options, based on both groups having a large sampling and clearly defined as belonging to humanities or math/science fields. Deardorff (2011, p. 69) noted the need for math and science

teachers to better understand and embrace concepts that integrate intercultural competencies into their curriculum. This was supported by Parkhouse (2016), who highlighted the dearth of research focused on the development of global competencies in math and science teachers. The researcher hypothesized that there would be a significant difference with teachers of modern foreign languages indicating a higher mean average than science and math educators. Prior to running the independent t-test, an examination of Levene's test demonstrated that assumption of homogeneity of variance was not met, p = .008, indicating heteroscedasticity. Therefore, Welshes t-test was used. Results indicated confirmed this hypothesis and indicated that the groups differed significantly on the overall assessment $t(255.318) = 9.552$, $p < .001$, $d = 1.167$. The means for teachers of modern world languages (M=4.288, SD=0.513) was significantly different from math and science educators (M=3.604, SD=0.636). Teachers of modern world languages had a significantly higher mean average than those who taught math and science (Table 16). This difference was significant across all factors and most apparent for the factor "curriculum" $t(255.651) = 12.56$, $p < .001$, $d = 1.517$. The means for teachers of modern world languages (M=4.350, SD=0.679) was significantly different from math and science educators (M=3.100, SD=0.920). These results support the construct validity of the ICME based on confirming the hypothesis based on the literature.

Table 16

Modern World Language (MWL) and Math and Science Teachers (MS)

Factor	*MWL (n=112)* *M/SD*	*MS (n=146)* *M*	*F*	*Sig.*	*t*	*df*	*p*	*d*
Curriculum	M=4.350 SD=0.679	M=3.100 SD=0.920	11.01	0.001	12.56	255.651	<.001	1.517
Student Inclusion	M=4.248 SD=0.585	M=3.700 SD=0.659	4.037	0.046	7.001	256	<.001	0.879
Cross-Cultural Openness	M=4.646 SD=0.450	M=4.259 SD=0.592	7.656	0.006	5.976	255.987	<.001	0.725
Collaboration and Adaptation	M=4.078 SD=0.664	M=3.297 SD=0.842	4.626	0.032	8.075	256	<.001	1.014
Systematic Awareness	M=4.114 SD=0.842	M=4.114 SD=0.699	2.687	0.102	5.464	256	<.001	0.686
Total Scale	M=4.288 SD=0.513	M=3.604 SD=0.636	7.196	0.008	9.552	255.318	<.001	1.167

Notes: Bonferroni correction was applied, setting the p-value for significance at <.008

Length of Time Living Outside of One's Primary Country

An ANOVA was used to compare the total mean scale score for the length of time they lived outside their primary country. Teachers who are often most successful at developing student global perspectives have had exposure to life outside of the mainstream (Siczek & Engle, 2017). Hammer (2011) noted the significance of living abroad, notably studying abroad, with increased intercultural competencies levels. Heinzmann et al. (2015) found that the length of time lived abroad corresponds to heightened levels of intercultural competence. However, stays of two months to a year can have a negative effect on intercultural competency levels due to encounters with everyday frustrations and challenges. The researcher hypothesized that responses on the ICME would indicate a statistically significant difference, with faculty who lived abroad longer having a higher mean average response based on having lived abroad and time lived abroad. To ensure an adequate sample size for each group, "Never" lived abroad was kept on its own (*n*=039); "> 3 months" lived abroad (*n*=58), "3-6 months" lived abroad (*n*=69),

and “7-11 months” lived abroad (*n*=18), were combined into “< 1 Year” lived abroad (*n*=145); “1-2 years” lived abroad (*n*=77), and “3-5 years” lived abroad (*n*=51), were combined into “< 1-5 Years” lived abroad (*n*=128); and “6 to 10 years” lived abroad (*n*=33), and “> 10 years” lived abroad (*n*=80), were combined into “> 5 Years” lived abroad (*n*=113). Prior to running the ANOVA, an examination of Levene’s test demonstrated that assumption of homogeneity of variance was met, p = .399. Results indicated that groups differed significantly on the overall assessment with $F(3.691) = 17.388$, $p < .001$, ηp2 = .07. This indicates that increased time lived abroad has a small effect (Tabachnick & Fidell, 2007, p.55). Groups differed significantly in all five factors (Table 17). Post-hoc analyses using Tukey's HSD test showed that participants in all groups differed significantly except for the “never” group and the “< 1year” group (p = 0.07). These findings support the hypothesis that participants who have spent more living abroad, particularly for one year or more, have a higher level of intercultural competence, supporting the ICME’s construct validity.

Table 17

ANOVA for Time Lived Abroad

Factor	Never *n*=309	< 1 year *n*=145	1-5 years *n*=128	> 5 years *n*=113	*df*	F	*p*	ηp2
Curriculum	M=3.461 SD=0.811	M=3.662 SD=0.891	M=3.891 SD=0.824	M=4.204 SD=0.743	3,691	25.411	< .001	0.10
Student Engagement	M=3.964 SD=0.593	M=3.910 SD=0.635	M=4.053 SD=0698	M=4.317 SD=0.524	3,691	11.377	< .001	0.05
Cross-Cultural Openness	M=4.416 SD=0.493	M=4.437 SD=0.555	M=4.500 SD=0.552	M=4.664 SD=0.412	3,691	7.035	< .001	0.03
Collaboration and Adaptation	M=3.590 SD=0.727	M=3.639 SD=0.795	M=3.827 SD=0.715	M=4.081 SD=0.645	3,691	14.1	< .001	0.06
Systematic Awareness	M=3.867 SD=0.647	M=3.869 SD=0.674	M=3.983 SD=0.689	M=4.186 SD=0.599	3,691	7.365	< .001	0.03
Full Scale	M=3.859 SD=0.557	M=3.903 SD=0.613	M=4.051 SD=0.615	M=4.290 SD=0.477	3,691	17.388	< .001	0.07

Notes:

Race and Ethnicity

An independent-samples t-test was conducted to compare responses from participants who identified as white with participants who identified as people of color. While demographic information was collected for multiple races and ethnicities, combining these groups in "people of color" allows a large enough sampling (n=124) to compare with a majority of participants (n=535) who identified as white. People of color must often learn to navigate cross-culturally between their own culture and that of a dominant culture. Malewski and Phillion (2009) observed the impact of racial identity and ethnicity in shaping how one interprets social norms and perceives being the "other." White teachers often view the world through the lens of social capital whereas non-white colleagues often navigate a society that is different from their own home culture. Zhao (2010) notes the limitation white teacher face working with students of different cultural backgrounds. The researcher hypothesized that there would be a significant difference with teachers who identify as people of color indicating a higher mean average than to white counterparts. Prior to running the independent t-test, an examination of Levene's test demonstrated that the assumption of homogeneity of variance was met, p = .126. Results confirmed this hypothesis and indicated that the groups differed significantly on the overall assessment $t(657) = 5.130$, p <.001, $d = .511$. The means for faculty who identified as people of color (M=4.201, SD=0.521) was significantly different from faculty who identified as white (M=3.917, SD=0.584). Faculty who identified as people of color have a significantly higher mean average than those who identified as white for all five factors (Table 18). This difference was most apparent for the factor "systematic awareness", $t(657) = 5.461$, p <.001, $d = .544$. The means for faculty who identified as people of color (M=4.224, SD=0.587) was significantly different from faculty who identified as white (M=3.876, SD=0.651). For this factor, faculty who

identified as people of color tended to be higher than those who identified as white. These results support the construct validity of the ICME based on confirming the hypothesis based on the literature.

Table 18

Race/Ethnicity (White and People of Color)

Factor	*POC (n=124)* *M/SD*	*White (n=535)* *M/SD*	*F*	*Sig.*	*t*	*df*	*p*	*d*
Curriculum	M=3.932 SD=0.811	M=3.641 SD=0.869	0.718	0.397	3.407	657	<.001	0.340
Student Inclusion	M=4.245 SD=0.595	M=3.972 SD=0.624	0.413	0.506	4.429	657	<.001	0.441
Cross-Cultural Openness	M=4.631 SD=0.406	M=4.444 SD=0.517	11.63	<.001	4.357	225.404	<.001	0.374
Collaboration and Adaptation	M=4.016 SD=0.677	M=3.651 SD=0.742	1.709	0.192	5.017	657	<.001	0.500
Systematic Awareness	M=4.224 SD=0.587	M=3.876 SD=0.651	1.934	0.165	5.461	657	<.001	0.544
Total Scale	M=4.201 SD=0.521	M=3.917 SD=0.584	2.342	0.126	5.130	657	<.001	0.511

Notes: POC = People of Color; Bonferroni correction was applied, setting the p-value for significance at <.008

Gender

An independent-samples t-test was conducted to compare responses from participants who identified as female with participants who identified as male. Malewski and Phillion (2009) gender shapes global perceptions and how one navigates cross-cultural settings. While Hammer (2011) found no significant difference due to gender on the IDI scale, Jankowski (2009) observed that males scored higher in the sub score of the "denial/defense" than female participants on the IDI. The researcher hypothesized that there would be a significant difference in the factor "Cross-Cultural Openness" with teachers who identified as "female" indicating a higher mean average than those who identified as "male." However, the researcher also hypothesized that there would not be a statistically significant difference on the overall scale.

Participants who identified as non-binary or third gender, *n*=2 (0.3), or "prefer not to identify," *n*=19 (2.7%), were not included due to the limited sample size. Prior to running the independent t-test, an examination of Levene's test demonstrated that assumption of homogeneity of variance was not met, p = .034, indicating heteroscedasticity. Therefore, Welshes t-test was used. Results confirmed both hypotheses while also noting a statistically significant difference for the factor "systematic awareness." For the overall scale, groups did not differ significantly $t(320.938) = -1.485$, $p < .069$, $d = -.113$ (Table 19). For the factor "Cross- Cultural Openness", the groups differed significantly on the overall assessment $t(675) = -3.698$, $p < .001$, $d = .339$. The means for faculty who identified as female (M=4.526, SD=0.479) was significantly different from faculty who identified as male (M=4.353, SD=0.577). Unexpectedly, for the factor "systematic awareness", the groups differed significantly on the overall assessment $t(675) = -2.417$, $p < .008$, $d = .219$. The means for faculty who identified as female (M=3.980, SD=0.627) was significantly different from faculty who identified as male (M=3.835, SD=0.738). Faculty who identified as female had a significantly higher mean average than those who identified as male for both factors. This supports construct validity, as hypothesized based on the literature, and opens the possibility for an increased understanding of the new factor "systematic awareness" as it relates to gender.

Table 19

Gender

Factor	Male (*n*=196) *M*	Female (*n*=481) *M*	*F*	*Sig.*	*t*	*df*	*p*	*d*
Curriculum	M=3.761 SD=0.894	M=3.671 SD=0.856	0.000	0.995	10227	675	0.110	0.104
Student Inclusion	M=3.960 SD=0.686	M=4.052 SD=0.601	3.201	0.074	-1.721	675	0.043	-0.146
Cross-Cultural Openness	M=4.353 SD=0.577	M=4.526 SD=0.479	8.775	0.003	-3.698	310.086	<.001	-0.339
Collaboration and Adaptation	M=3.668 SD=0.795	M=3.743 SD=0.731	2.363	0.125	-1.175	675	0.120	-0.100
Systematic Awareness	M=3.835 SD=0.738	M=3.980 SD=0.627	8.105	0.005	-2.417	315.252	0.008	-0.219
Total Scale	M=3.916 SD=0.649	M=3.994 SD=0.563	4.497	0.034	-1.485	320.938	0.069	-0.133

Notes: Bonferroni correction was applied, setting the p-value for significance at <.008

Additional Measures

An ANOVA was run to compare the number of years participants indicated that they had been in an education career. The findings were insignificant for all five factors and the total scale score $F(5.689) = 1.340$, p = 0.246, n2p = .01. Similarly, an ANOVA was run to compare grade level taught and found to be statistically insignificant $F(3.691) = 0.561$, p = 0.641, n2p = .002. An independent t-test was run to compare responses from participants who taught in Canada with participants who taught in the United States and found to be statistically insignificant $t(693) = 0.432$, p = .333, $d = 0.041$. No further inquiries were conducted. These results suggest a lack of significant differences in categories that would not typically be associated with intercultural competencies, which further supports the construct validity of the ICME.

Summary

The development of the ICME began with a review of the literature and contributions from a panel of experts. Following data collection from *n*=695 in-service educators at *n*=16 schools in the United States and Canada, an exploratory factor analysis challenged the four-

factor theoretical model. It provided statistical support for a five-factor model. This was further refined following a confirmatory factor analysis and consideration of modification recommendations. The result was a five-factor model that demonstrated a strong model fit and was logically sound.

Demographic characteristics commonly associated with contributing to intercultural competencies were examined to see if they predicted the results of the ICME based on the new five-factor theoretical model. Five demographic characteristics associated with intercultural competencies, including the ability to speak a second language, experience living outside one's primary country during formative years, length of time lived outside of one's primary country, race/ethnicity, and subject area taught, were all found to indicate significant differences in assessment outcomes. Three demographic characteristics less associated with intercultural competencies, such as gender, grade level taught, and years in education, were all found to indicate insignificant differences in assessment outcomes. The observed variance in demographic outcomes validates initial assumptions and demonstrates that the ICME measures the intended latent variable of intercultural competencies. The results of this study will be contextualized and further interpreted in the next chapter in order to draw pivotal scholarly discussions and implications for practices and policy.

CHAPTER 5: DISCUSSION

The purpose of this study was to develop and validate the Intercultural Competency Measure for Educators (ICME), an intercultural competencies assessment instrument for K-12 in-service educators. Given the lack of a previous instrument targeting this specific demographic, the results of this study benefit both practitioners and researchers in providing an empirically validated framework for educators' intercultural competencies. This has the potential to promote critical reflection, enhance professional development, and measure the effect of targeted interventions.

Three research questions guided this study:

1. *What is the extent to which content validity is established in the design of the Intercultural Competency Measure for Educators (ICME)?*
2. *To what extent does the ICME reveal construct validity evidence for its intended use?*
3. *To what extent do demographic characteristics commonly associated with contributing to intercultural competencies growth predict results of the ICME based on the Five-factor theoretical model?*

The first four chapters introduced the background of this study, reviewed related literature, explained the research procedures, and examined the findings. These chapters provide insight into the original four-factor theoretical model, which served as the foundation for this study. This theoretical model was modified in Chapter Four to include five factors and is discussed in this chapter's findings and implications.

This study used a combination of expert input and quantitative data to examine the construct validity of a measure of intercultural competencies among in-service educators teaching at schools that prioritize global education. The methodology and findings chapters

describe how an Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were used to examine construct validity. Finally, parametric tests, including independent-t tests and one-way ANOVA were used to explore demographic data to further support construct validity.

This chapter provides interpretations of the findings by drawing connections to the literature while addressing each of the three research questions. The insight provided by these interpretations then fosters the exploration of implications for practitioners and researchers. This chapter concludes with a review of study limitations offering recommendations for future research.

Research Question One - Findings

The development of the ICME required establishing content validity through examining previous intercultural competency measures, reviewing the literature, and consultation with a panel of experts. Prior researchers have challenged the premise that intercultural competencies can be measured quantitatively (Byram, 2021; Álvarez Valdivia & González Montoto, 2018; Deardorff, 2016, p. 122). Widely used intercultural competencies assessment instruments, including the Intercultural Development Inventory (IDI), were not designed to be specifically applicable to K-12 educational settings (He et al., 2017; Bustamante et al., 2017). The Teacher Global Readiness (TGR) scale was developed for educators but targeted those who teach literature (Kerkhoff, 2017). Deardorff (2011) noted that nurturing intercultural competencies is essential across the curriculum and all grade levels. This belief guided the development of the scale examined in this study.

This study was strongly influenced by the theoretical framework of Critical Cosmopolitanism, which embodies heightened levels of consciousness and reflection while

enhancing relationships between globally diverse people (Byker, 2016; Kerkhoff, 2017; Müller et al., 2020; Oxley & Morris, 2013). Critical Cosmopolitanism aligns with the increasing emphasis schools place on intercultural competencies development related to the attitudes, skills, and knowledge necessary for navigating, engaging in, and impacting an increasingly globalized society (Byram, 2021; Deardorff, 2006; Kerkhoff, 2017).

Regular feedback from a panel of experts complemented findings from the literature and prior intercultural competencies instruments to develop a measure relevant to modern K-12 in-service educators. This panel of experts provided valuable insight based on experience in the classroom, observations of teachers, knowledge of past literature, and understanding of methodological approaches. This diverse range of perspectives helped guide each stage of the scale design and interpretation of data from the EFA and CFA (Carpenter, 2007). The panel of experts helped ensure the measure remained applicable to the needs of in-service educators, diverging from the heavy focus on pre-service teachers in the literature.

The understanding gained from the literature, panel of experts, and prior assessment instruments helped to guide the development of a pilot study. This pilot study provided data from a small sample (n=97) informing the creation of the initial four-factor theoretical model that included: *Intercultural Curriculum Content, Intercultural Learning Environment, Student Engagement,* and *Personal Intercultural Competencies Growth*. Key themes from the literature and feedback from an expert panel informed this model. Recommendations from the panel of experts, in addition to further insight from the literature, later helped contribute to the identification of the five factors that emerged from the EFA and CFA, further establishing content validity by ensuring decisions combined statistical findings with practical insight.

Research Question Two - Findings

The researcher examined the construct validity for the four-factor theoretical model using EFA and CFA results based on data collected from the ICME. A review of the EFA factor loadings revealed the potential for a five-factor model and challenged the placement of items in the four-factor theoretical model. The CFA confirmed a better model fit for the five-factor model and provided modification recommendations that were considered alongside a review of the literature and feedback from the panel of experts. The new five-factor model included the latent constructs of (1) *Curriculum;* (2) *Diverse Student Inclusion*; (3) *Cross-Cultural Openness*; (4) *Collaboration and Adaptation*, and (5) *Systematic Awareness*. A notable difference between the four-factor theoretical model and the five-factor model was the adaptation of the theoretical factor of *Intercultural Learning Environment* into two separate factors, *Collaboration and Adaptation* and *Systematic Awareness*. This new model provided a framework that is more inclusive of the whole school environment, how teachers work together, and the relationship between educators and the context in which they teach.

Factor One: Curriculum

The first factor is in line with the initial theoretical framework and the trend of schools moving towards incorporating a focus on global and intercultural topics throughout their curriculum. While modern world language and humanities curricula have traditionally taken steps to address this content, educators in all subject areas need to be encouraged to link content to diverse cultures and perspectives (Byram, 2021, p. 122; Colvin & Edwards, 2018; Kerkhoff, 2017; Putman & Byker, 2020). Smolcic and Katunich (2017) noted that a cross-curricular approach to intercultural competencies development is vital as society becomes more global. Items in this factor assess the value educators place on second language proficiency while also

targeting how global content is embedded and evaluated. The teacher's understanding of global perspectives and how this is reflected in the curriculum is also examined. As teachers develop a clearer understanding of the cultures represented by their students, they are better able to adjust their cultural lens and reflect critically on how appropriate curriculum material is selected (Biraimak & Jotia, 2013; Heizmann et al., 2015; van Werven et al., 2023). This enables students to better see themselves in the curriculum while simultaneously being exposed to the diverse background of their peers.

Factor Two: Diverse Student Inclusion

The second factor blends two factors from the original theoretical model, *Student Engagement* and aspects of the *Learning Environment*, to focus on how educators promote the inclusion of globally diverse students. Byram (1997) notes that preparing for increased levels of global diversity in the classroom is a cross-curricular priority. In addition to serving the needs of engaging students of diverse backgrounds, this response to cultural norms is essential when preparing all students to be globally responsible citizens (Guo, 2014). This factor addresses how the development of intercultural competencies expands beyond the curriculum by promoting awareness of others, empathy, and adaptability (Biautti, 2018; Morais & Ogden, 2011). Putnam and Byker (2020) highlight that as teachers develop a more empathetic mindset, they are better prepared to establish a learning environment inclusive of all students lived experiences. The items in this factor link to teachers' actions to ensure that globally diverse students feel valued and have a voice. The theme of advocacy also links these items, promoting the role of educators in ensuring that students of marginalized backgrounds are recognized, supported, and integrated into classroom settings.

Factor Three: Cross-Cultural Openness

The third factor is consistent with the original theoretical model's factor of *Cross-Cultural Openness* and is primarily focused on teacher attitudes and values. The process of becoming more interculturally competent often involves making sense of the diversity of our world through actively observing, listening, and understanding different perspectives (Hauserwas et al., 2021; Kerkhoff, 2017; Wahlström, 2014). Byker and Putnam (2019, p. 88) note that this openness to exploring new cultures is essential to expanding one's level of consciousness. Becoming more aware of diverse cultural backgrounds often sparks self-reflection, promoting consciousness of one's ethnocentrism while fueling a desire to pursue positive cross-cultural interactions (Bernardes et al., 2019; Kerkhoff & Cloud, 2020; Martorana et al., 2021). Items in this factor embody the importance of educators remaining lifelong learners as they harness their curiosity to explore diverse cultures represented by their students and those in the broader community. This expanded consciousness and awareness have the potential to shape how educators interact with culturally diverse students while ensuring all students are exposed to a more globally inclusive learning experience.

Factor Four: Collaboration and Adaptation

The fourth factor builds upon the original theoretical model by focusing on strategies and collaborative approaches that target the implementation of a global curriculum in a culturally inclusive learning environment. Collaboration between educators helps to infuse school-wide approaches that promote global curriculum while incorporating students' learning and emotional needs (Ghosn, 2020; He et al., 2017; Kerkhoff & Cloud, 2020). Items in this factor include discussing strategies between colleagues and a collective approach toward promoting a more culturally inclusive school. Educators also have opportunities to strategize within their classroom

environment by promoting intercultural dialogue and understanding, opening the door for greater respect for cultural diversity, awareness of different perspectives, and engagement in reflection on the student's role as a citizen of the world (Byker, 2016; Kerhoff, 2017; Müller et al., 2020; Pashby et al., 2020; van Werven et al., 2023). Items in this factor emphasize the importance of adaptability regarding how educators approach global curriculum and how they engage with people of diverse backgrounds in a school environment.

Factor Five: Systematic Awareness

The fifth factor also builds upon the original theoretical model by emphasizing awareness of systemic attributes that can challenge or influence the engagement, inclusion, or learning experience of culturally diverse students. Quite often, inequities and power dynamics can permeate environments where cultural and language differences are not addressed (Deardorff, 2006; Kerkhoff & Cloud, 2020; Larsen & Searle, 2017). Educators have opportunities to transcend beyond the neoliberal emphasis on developing global competitiveness by adapting an approach that addresses attitudes and perceptions of inequities while embracing opportunities for critical reflection and recognition of power dynamics (Bernardes et al., 2019; Choi & Shin, 2016; O'Connor & Zeichner, 2011). Kerkhoff and Cloud (2020) note the need for greater understanding that leads to transformational change, better-enabling educators to address inequalities related to power, privilege, and oppression. While most items in this factor focus on the acknowledgement and assessment of inequalities and power structures, there is also a link to the need for educators to act by breaking down stereotypes and having a positive impact on their learning environment.

Factor Comparisons

Items representing the four intercultural competencies domains of *attitudes, skills, knowledge*, and *actions*, which were evenly represented in each of the factors of the theoretical model, were more clustered in the five-factor model. This domain of *attitude* is represented by all five items in the factor of *cross-cultural openness.* Deardorff (2011) notes the importance of *attitude* in relating to attributes of openness and curiosity, both of which are central to the theme of this factor. The domain of *action* is highly represented in the factors of *diverse student inclusion,* and *collaboration and adaptation*. These factors embrace a teacher's capacity to promote change, representing actions that impact students and an ever-changing school environment. The domain of *knowledge* is most prevalent in items placed in the new factor of *systematic awareness*. Martorana et al. (2021) note the significance of knowledge in promoting awareness of inequities and power structures. Only the domain of *skills*, which Deardorff (2011) links to *listening, observing, evaluating, interpreting*, and *reflecting*, is represented by items in four factors, signifying the importance of these skills in all areas of intercultural competencies development. While the four domains are not evenly spread amongst the five factors, all are well represented in the overall model.

When examining the correlations between five factors following the CFA, concern exists that divergent validity is violated due to high correlations between the factor *Collaboration and Adaptation* and other factors, including Curriculum (0.904), Systematic Awareness (0.884), and Diverse Student Inclusion (1.004). High correlations between factors signal a risk of overlap where different factors might measure the same characteristics. When examining effect size using parametric tests, differences emerged between factors with high correlations. The notable exception was between student inclusion and correlation and adaptation. The statistical

similarities between these factors did not overrule the practical distinctions noted by the researcher and panel of experts. The relationship between highly correlating factors could be more closely examined in future research, with additional items added to each factor to help clarify distinctions.

Research Question Three - Findings

To further explore construct validity, the third research question focused on the relationship between results from items on the ICME and demographic characteristics known to be associated with intercultural competencies. Demographic variables examined included *the ability to speak a second language*, *race/ethnicity* (white and people of color), *subject areas taught*, *number of years teaching*, and *grade level taught*. Parametric tests, including an independent t-test and one-way ANOVA were compared to hypotheses drawn from the literature to see if they were consistent with results from the ICME.

The parametric test results generally aligned with the hypothesis for each demographic variable examined. For demographic categories, including the *ability to speak a second language*, *subject area taught* (modern world language and science/math), and *race and ethnicity* (people of color and white), an independent t-test supported their respective hypotheses and was consistent across all five factors and the overall model. The means for people who spoke a second language was consistently higher than those who were monolingual, particularly for the factor of *curriculum.* The same pattern was evident for the *ability to speak a second language*, and *subject area taught* (modern world language and science/math), with teachers of modern world language having a consistently higher mean than those who taught math/science, particularly for the factor of *curriculum*. It is likely that those who speak multiple languages are more predisposed to seeing the importance of blending intercultural competencies into the

curriculum. The means for people who identified as people of color was consistently higher than those who identified as white, particularly for *systematic awareness.* This is consistent with past findings that people of color are more aware of systematic inequality due to their experience being the "other" and having to adapt to a majority culture that is different from their own (Malewski & Phillion, 2009; Zhao, 2010).

The independent t-test results for *gender* only indicated significant results for *cross-cultural openness* and *systematic awareness*. For both factors, female educators had a higher means average than male educators. This was consistent with the finding of Jankowski (2009), who found that males had a higher mean for the sub-score "denial/defense" than females. Male participants may be more protective of their own beliefs rather than being open to the cultures and perspectives of others. The results for *systematic awareness* were surprising and supported the need for further research into this new factor that emerged from the study.

The one-way ANOVA results for *time lived abroad* supported the hypothesis that those who lived abroad longer would have a greater mean average than those with less than a year or who had never done so. These results support prior research noting that more time spent living abroad results in higher levels of intercultural competence (Hammer, 2011; Heinzmann et al., 2015). The most significant effect of time spent abroad was on the factors of *curriculum,* and *collaboration and adaptation*. This result for *curriculum* is consistent with past observations that those who spent time abroad were more successful at developing students' global perspectives (Siczek & Engle, 2017). The comparatively higher effect of experience lived abroad on *collaboration and adaptation* provides evidence supporting this distinct factor that exhibited high correlations with other factors in the CFA. The post hoc analysis of the ANOVA indicated that the mean difference between those who never lived abroad and those who lived abroad for

one year or less was insignificant. In contrast, lengthier time spent abroad had a significant mean difference. This challenges Heinzmann et al. (2015) findings that the experience of living abroad has a significant effect during the first six weeks and tends to plateau after more extended periods. Collectively, the results of the parametric tests examining the demographic variables often associated with intercultural competencies, supported the construct validity of the ICME and provides unique opportunities for further research.

Implications for Practitioners

Rapid developments in our modern world have fueled the demand for schools to better address the need to prepare students for participation in an increasingly interconnected global society (Martorana et al., 2021; OECD, 2018). While practical considerations often promote the neoliberal approach to global education, this study embraced critical cosmopolitanism as an outlook for educators seeking to enable students to appreciate diverse perspectives, contemplate the interdependence between people, and respond to existing inequalities (Byker, 2016). This study provides an empirically validated measure for assessing educators' intercultural competencies based on the critical cosmopolitanism framework designed specifically for K-12 learning environments. In addition to addressing the need to embrace intercultural competencies in the curriculum, this measure provides insight into how educators interact with globally diverse students, develop their global perspectives, collaborate with colleagues, and interact with the systems and context in which they teach.

By providing a quantitative approach for assessing educators' intercultural competencies, school leaders are better able to align professional development with specific growth needs and further understand the effect of interventions once delivered. A theoretically grounded scale has the potential to provide the data needed to inform how financial support and human resources are

devoted (Kerkhoff, 2017; Morais & Ogden, 2011). Fantini (2009) notes the importance of transparency concerning how the intercultural competencies being measured connect with educational objectives. Rather than addressing educators' intercultural competencies in their entirety, this model provides five distinct factors that allow for a more strategic and purposeful approach to professional development. The ICME's quantitative approach allows schools to collect data more efficiently, allowing for the longitudinal effect of professional interventions to be chronicled and evaluated (Baiutti, 2018; Jankowski,2019; Kerkhoff, 2017). The use of this quantitative assessment instrument, in combination with existing qualitative assessment techniques, allows for a mixed-method approach that enable teachers and school leaders to better reflect on the progression of intercultural competencies development (He et al., 2017). The combined data from the mixed-methods approach will also provide educators with more opportunities for critical reflection as they progress in areas related to each of the five factors.

While the impact of a more globalized society is evident, it is often vague how educators best address the need for intercultural competencies development in learning environments. The five-factor model validated in this study provides educators clarity in distinct areas that impact them professionally, along with a better understanding of what is important. The insight provided for an effectively delivered assessment instrument has the potential to promote ongoing critical reflection while guiding professional development. Ultimately, students will benefit from the influence of educators who are more purposeful in how they model intercultural competencies while incorporating related themes, like advocacy, into the relationship they build and overall pedagogy.

Two factors that emerged from this study, *collaboration and adaptation* and *systematic awareness*, were less prevalent in prior literature yet are of noted importance to in-service

practitioners. As the development of intercultural competencies is applied across all subject areas, faculty *collaboration and adaptation* allow for the sharing of resources and strategies. Technology has also increased the ability of teachers to collaborate not just within-schools but also between-schools as new approaches for global curriculum growth and diverse student inclusion are explored (Byker, 2016; Kerkhoff, 2017). *Systematic awareness* acknowledges how systematic inequalities, power structures, and rigid hierarchies have marginalized culturally diverse individuals. Recent societal movements, including #MeToo and #BlackLivesMatter, have impacted educational settings and how educators address often unspoken barriers (Arya, 2022). Both *collaboration and adaptation* and *systematic awareness* address the call to action, a central component of critical cosmopolitanism (Byker & Putnam, 2019). Including these factors in the educators' intercultural competencies ensure that teachers not only embed related themes in their curriculum but also model active engagement as global citizens aiming to build a more equitable and purposeful learning environment for all.

Implication for Researchers

The increased prioritization of global education and the desire to address expanding global diversity in schools has promoted research into educator intercultural competencies development and assessment. Utilizing current literature, contributions from a broad panel of experts, and data from a pilot examination, this study established content validity for an assessment instrument based on a theoretical four-factor model. Data were collected from a large sampling of in-service K-12 educators from 16 schools in two countries, which was used to run an EFA and CFA, further examining the four-factor theoretical model, and providing support for a refined five-factor model. Findings from the parametric tests and feedback from the panel of experts further supported the construct validity of this new model. The use of a large and diverse

sampling also responded to the need expressed by prior researchers for a greater understanding of construct validity related to how intercultural competence is defined (Hammer, 2011; Jankowski, 2019; Kerkhoff, 2017). In addition to providing more representative and generalizable data, this study examines the means differences in subgroups to test hypotheses and establish construct validity. These parametric findings supported prior research while establishing opportunities for further exploration into how subgroups relate to specific factors in the new model.

The population surveyed in this study provides unique and more generalizable insight for researchers into the intercultural competencies of educators. No known prior study examines K-12 educator intercultural competencies across all grade levels and subject areas. Deardorff (2011) notes the importance of the universal application of intercultural competencies skills in all academic areas. The ICME examined in this study avoids content-specific and grade-specific items, allowing a holistic approach to measuring educator intercultural competencies. While the IDI is the most broadly used intercultural assessment instrument, the ICME offers researchers an empirically validated alternative designed to be specifically applicable to schools. This is useful to researchers in ensuring the specificity, reliability, and relevance of the data collected.

This study highlights the relevance of educator intercultural competencies beyond the impact on the *curriculum* and *student engagement*. While some research has examined *cross-cultural awareness*, *collaboration* and *systematic awareness*, these areas have yet to be included in a prior intercultural competencies assessment instrument or quantitatively measured for the impact on the role of a teacher in school environments. This study provides evidence that these factors are relevant and central to understanding how educator intercultural competencies are

nurtured and assessed. Researchers have opportunities to further research these new factors and how they relate to the needs of schools adapting to a more globalized society.

Limitations of The Study

Several limitations to this study impact reliability and validity. First, the data gathered from K-12 educators were entirely self-reported, potentially contributing to construct-irrelevant variance related to response bias and possible misinterpretation of items. Self-reported data were noted limitations of other intercultural competencies assessment instruments (Alvarez Valdivia & Gonzalez Montoto, 2018; Kerkhoff, 2017; Morais & Ogden, 2011). Participation in this study was optional and it was possible that educators who chose to participate already had a predisposition to global education. Since the ICME was presented to educators by school administrators at various participating schools, there was limited control over how the ICME was delivered. Some schools sent out the ICME by electronic mail, being completed during or after school hours. In contrast, other schools provided time for completion during set times at faculty meetings.

Participants for this study were selected using convenience sampling with no screening procedure other than being current educators at a GEBG member school or international school (Huck, 2004, p. 111). While this approach allowed for contextual similarities between participants at the schools in this study, it likely inhibited external validity and the generalizability of the findings. The possibility exists that results would be different if data were gathered from faculty at public schools, particularly those serving marginalized populations. Bernardes et al. (2019) note the need for greater research on intercultural competencies growth at low-income schools. Public schools are less likely to have fully developed global education programs, and state policy often dictates what emphasis, if any, is placed on intercultural

competencies development. The applicability of this study to educators at public schools provides opportunities for future research studies.

The ICME was only offered in English, limiting the ability to collect data from non- or limited-English speakers and potentially weakening construct validity. Deardorff (2016, p. 96) notes that using multiple languages in an assessment helps to ensure representation and avoid "cultural and linguistic domination." One advantage of broadly used IDI is its availability in multiple languages (Hammer, 2011). The application of multiple languages to future studies of the ICME will benefit construct validity and minimize construct-irrelevant variance by better ensuring the interpretation of items is correct and responses are measuring the desired construct.

The lack of a comparable scale explicitly addressing the assessment of K-12 educators' intercultural competence in all subject areas limited the ability to check for the concurrent validity of the ICME. Concurrent validity helps examine how well the outcome of one instrument correlates with results from previously validated measures using similar constructs (Huck, 2004, p. 90). The results of this study could be beneficial in providing opportunities for checking concurrent validity for future K-12 educators' intercultural competencies scales.

Recommendations for Future Research

In addressing several of the previously mentioned limitations, it is recommended that future researchers offer translated versions of the ICME, allowing participants to complete the measure in their primary language. This would help minimize construct-irrelevant variance, strengthen external validity, and allow for a broader range of culturally diverse participant participation.

This study was limited to in-service educators at independent schools in the United States and Canada, all of which are a common network of global education schools. Future research

could incorporate data from a broader range of public and independent schools to increase reliability and construct validity. This could also increase the generalizability of findings and conclusions, allowing schools with the greatest need to benefit from recommendations. Future research could also be expanded to include data from schools in diverse global regions. In addition to providing insight into reliability and construct validity, this data could help examine how native culture and context effects educators' intercultural competencies measures.

This study examines an empirically validated measurement instrument providing a quantitative approach to understanding longitudinal educators' intercultural competencies development. Prior researchers have called for a better understanding of the longitudinal effect of intercultural competencies interventions (Jankowski,2019; Kerkhoff, 2017; Okken et al., 2019). As additional items are created to measure each of the five factors, future researchers will have opportunities to investigate the longitudinal effects of interventions, including professional development, both domestically and abroad. This insight has the potential to guide improvements to these programs while helping to inform schools on how best to apply limited time and financial resources.

It is recommended that future researchers utilize multiple regression to further understand the relationship of between various demographic variables and scores on the educators' intercultural competencies measure. A multiple regression study would allow for the evaluation of the significance of targeted demographic variables by controlling for the effect of other demographic variables. This data would help to further explore construct validity.

Researchers wishing to examine educator intercultural competencies further can better control for construct irrelevant variance by utilizing this validated quantitative instrument alongside other multiple measures of the same construct. This addresses a need for multiple

forms of assessment for educator intercultural competencies based on standardized, validated measures, allowing for greater concurrent validity (Jankowski, 2019; Kerkhoff, 2017). This allows for deeper insight into convergent validity, allowing for comparison with the results of future quantitative measurement instruments that measure similar constructs. An increased understanding of concurrent validity will provide a path for a deeper understanding of existing qualitative research, especially when the two approaches are used side-by-side with the same participants. This study further opens the doors towards expanding approaches towards defining and measuring educator intercultural competencies while allowing for a deeper understanding of related constructs, adding credibility to this field of research, and ultimately providing more reliable recommendations to practitioners.

Summary

This study presents the ICME, a reliable and validated K-12 in-service educators' intercultural competencies assessment instrument based on findings from a large sampling of in-service practitioners. Two new factors have emerged, *systematic awareness* and *collaboration and adaptation*. These factors reflect current societal demands and address the need for regular evaluation of how intercultural competencies are defined in school environments (Deardorff, 2006). Byram (2021), who is critical of quantifiable measures for educators' intercultural competencies, notes the potential benefits of a psychometrically sound scale in serving the needs of schools. The validated scale presented in this study provides researchers and school leaders with an empirically validated instrument for a more generalizable understanding of the effect of interventions while promoting opportunities for critical reflection and professional growth. In addition, this scale promotes a shift in the direction of global education in schools from neoliberal practicality to a more cosmopolitan perspective that encourages the enhancement of

intercultural relationships and actions that create a more equitable and inclusive environment. This scale provides a stepping-stone for future researchers wishing to build on similar quantitative measures while promoting the growth of educators' intercultural competencies as a means of preparing students for an increasingly interconnected global society.

REFERENCES

Alvarez Valdivia, I. M., & Gonzalez Montoto, I. (2018). Teachers' intercultural competence: A requirement or an option in a culturally diverse classroom? *International Journal of Inclusive Education, 22*, 510-526. https://doi.org/10.1080/13603116.2017.1377298.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

Arya, D. J. (2022). Into the void of discourse. *Linguistics and Education, 68*. https://doi.org/10.1016/j.linged.2021.100964

Baecher, L., & Chung, S. (2020) Transformative professional development for in-service teachers through international service learning. *Teacher Development, 24*(1), 33-51. https://doi.org/10.1080/13664530.2019.1682033

Bandalos, D. L., & Finney, S. J. (2019). Factor analysis. Exploratory and confirmatory. In G. Hancock & R. O. Mueller, (Eds.), *The reviewer's guide to quantitative methods in the social science* (2nd ed.) (pp. 99–122). Routledge.

Bernardes, R., Black, G., Jowi, J., & Wilcox, K. (2019). Teachers' critical interculturality understandings after an international teaching practicum. *Critical Studies in Education*, 1–17. https://doi.org/10.1080/17508487.2019.1620817

Baiutti, M. (2018). Fostering assessment of student mobility in secondary schools: Indicators of intercultural competence. *Intercultural Education* (London, England)*, 29*(5-6), 549–570. https://doi.org/10.1080/14675986.2018.1495318

Biraimah, K., & Jotia, A. (2013). The longitudinal effects of study abroad programs on teachers' content knowledge and perspectives: Fulbright-Hays group projects abroad in Botswana and Southeast Asia. *Journal of Studies in International Education, 17*(4), 433–454. https://doi.org/10.1177/1028315312464378

Bustamante, Skidmore, S. T., Nelson, J. A., & Jones, B. E. (2016). Evaluation of a cultural competence assessment for preservice teachers. *The Teacher Educator, 51*(4), 297–313. https://doi.org/10.1080/08878730.2016.1186767

Byker, E. J. (2016). Developing global citizenship consciousness: Case studies of critical cosmopolitan theory. *Journal of Research in Curriculum & Instruction, 20*(3), 264-275. https://doi.org/10.24231/rici.2016.20.3.264

Byker, E. J. & Ezelle-Thomas, V. (2021). Preparing teacher candidates with global competencies: Taking action on the global water crisis with service learning. *Journal of Research in Childhood Education, 35*(2), 268–280. https://doi.org/10.1080/02568543.2021.1880996

Byker, E. J., & Putman, S. M. (2019). Catalyzing cultural and global competencies: Engaging preservice teachers in study abroad to expand the agency of citizenship. *Journal of Studies in International Education, 23*(1), 84–105. https://doi.org/10.1177/1028315318814559

Byram, M. (1997) *Teaching and assessing intercultural communicative competence*. Multilingual Matters.

Byram, M. (2021). *Teaching and assessing intercultural communicative competence*: *Revisited*. Multilingual Matters. https://doi.org/10.21832/Byram0244

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* Routledge.

Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures, 12*(1), 25–44. https://doi.org/10.1080/19312458.2017.1396583

Cattell, R. B. (1966). The Scree Test for the number of factors. *Multivariate Behavioral Research, 1*(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Chandir, H., & Gorur, R. (2021). Unsustainable measures? Assessing global competence in PISA 2018. *Education Policy Analysis Archives*, *29*, (122), 1-27. https://doi.org/10.14507/epaa.29.4716

Cho, M.-H., & Cho, Y. (2017). Self-regulation in three types of online interaction: A scale development. *Distance Education, 38*(1), 70–83. https://doi.org/10.1080/01587919.2017.1299563

Choi, Y., & Shin, E. (2016). Rethinking the hermit kingdom: US social studies teachers' cross-cultural professional development in South Korea. *Asia-Pacific Journal of Teacher Education*, *44*(2), 188–202. https://doi.org/10.1080/1359866X.2015.1066490

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Lawrence Erlbaum Associates Publishers.

Colvin, R. L., & Edwards, V. (2018). *Teaching for global competence in a rapidly changing world*. OECD Publishing. https://doi.org/10.1787/9789264289024-en

Comrey, A. L. & Lee H. B. (1992). *A first course in factor analysis*. Lawrence Erlbaum.

Costello, A. B., & Osborne, J. (2005). Practical assessment, research & evaluation. *The Journal of Consumer Marketing, 10*(7), 1-9. https://doi.org/10.7275/jyj1-4868

Das, R., Risner, M., & Monroe, M. (2016). Infusing global competence in classrooms through short term teacher professional development abroad. *International Journal of Global Education, 5*(2), 28-45.

Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *International Journal of Market Research*, *50*(1), 61–104. https://doi.org/10.1177/147078530805000106

Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education, 10*(3), 241–266. https://doi.org/10.1177/1028315306287002

Deardorff. (2011). Assessing intercultural competence. *New Directions for Institutional Research*, *2011*(149), 65–79. https://doi.org/10.1002/ir.381

Deardorff, D. K. (2016). "How to assess intercultural competence." In Z. Hua (Ed.), *Research methods in intercultural communication: A practical guide* (pp.120–134). John Wiley & Sons.

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's *t*-test instead of Student's *t*-test. *International Review of Social Psychology*, *30*(1), 92–101. http://doi.org/10.5334/irsp.82

Downey, R. G., & King, C. V. (1998). Missing data in Likert ratings: A comparison of replacement methods. *The Journal of General Psychology, 125*(2), 175–191. https://doi.org/10.1080/00221309809595542

Driscoll, P., Rowe, J.E., & Thomae, M. (2014) The sustainable impact of a short comparative teaching placement abroad on primary school language teachers' professional, linguistic and cultural skills. *The Language Learning Journal, 42*(3), 307-320. https://doi.org/10.1080/09571736.2014.917332

Fantini, A. (2009). Assessing intercultural competence: Issues and tools. In D. K. Deardorff (Ed.), The Sage handbook of intercultural competence (pp. 456–476). Sage.

Fowler, F. J. (2014). *Survey research methods*. Sage.

Gallagher M.W. & Brown T.A. (2013). Introduction to confirmatory factor analysis and structural equation modeling. In T. Teo (Ed.), *Handbook of quantitative methods for educational research* (pp, 289–314). Sense Publishers. https://doi.org/10.1007/978-94-6209-404-8_14.

Gastwirth, J. L., Gel, Y. R., & Miao, W. (2009). The impact of Levene's Test of Equality of Variances on statistical theory and practice. *Statistical Science*, *24*(3), 343–360. https://doi.org/10.1214/09-STS301

Global Education Benchmark Group (2023, January 5). https://gebg.org/.

Goren, H., & Yemini, M. (2017). The global citizenship education gap: Teacher perceptions of the relationship between global citizenship education and students' socio-economic status. *Teaching and Teacher Education*, *67*, 9–22. https://doi.org/10.1016/j.tate.2017.05.009

Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in Nursing & Health, 20*(3), 269–274. https://doi.org/10.1002/(SICI)1098-240X(199706)20:3<269::AID-NUR9>3.3.CO2-3.

Guo, L. (2014). Preparing teachers to educate for 21st century global citizenship: Envisioning and enacting. *Journal of Global Citizenship & Equity Education, 4*(1), 1–23.

Hammer, M.R. (2011). Additional cross-cultural validity testing of the Intercultural Development Inventory. *International Journal of Intercultural Relations, 34*(4), 474–487. https://doi.org/10.1016/j.ijintrel.2011.02.014

Hammer, M.R. (2012). The Intercultural Development Inventory: A new frontier in assessment and development of intercultural competence. In M. Vande Berg, R. M. Paige, & K. H. Lou (Eds.), *Student Learning Abroad* (pp. 115–136). Stylus Publishing.

Hauerwas L.B., Kerkhoff, S. N., & Schneider, S. B. (2021). Glocality, reflexivity, interculturality, and worldmaking: a framework for critical global teaching. *Journal of Research in Childhood Education, 35*(2), 185–199. https://doi.org/10.1080/02568543.2021.1900714

He, Y., Lundgren, K., & Pynes, P. (2017). Impact of short-term study abroad program: Inservice teachers' development of intercultural competence and pedagogical beliefs. *Teaching and Teacher Education*, *66*, 147–157. https://doi.org/10.1016/j.tate.2017.04.012

Heinzmann, S., Künzle, R., Schallhart, N., & Müller, M. (2015). The effect of study abroad on intercultural competence: Results from a longitudinal quasi-experimental study. *Frontiers 26*(1), 187–208. https://doi.org/10.36366/frontiers.v26i1.366

Huck, S. W. (2004). *Reading statistics and research*. Pearson Educations Inc.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hur, J. W., & Suh, S. (2018). Community-university partnerships in practice: Development of welcoming learning environments for new immigrants. *Journal of Higher Education Outreach and Engagement*, *22*(3), 45–67.

Jankowski, J.P. (2019) A construct validation argument for the Intercultural Development Inventory. *Measurement and Evaluation in Counseling and Development, 52*(2), 75-89, https://doi.org/10.1080/07481756.2018.1497428

Kambutu, J., & Nganga, L. W. (2008). In these uncertain times: Educators build cultural awareness through planned international experiences. *Teaching and Teacher Education, 24*, 939-951. http://dx.doi.org/10.1016/j.tate.2007.08.008

Kerkhoff, S. (2017). Designing global futures: A mixed methods study to develop and validate the teaching for global readiness scale. *Teaching and Teacher Education, 65*, 91–106. https://doi.org/10.1016/j.tate.2017.03.011

Kerkhoff S.N., & Cloud, M. E. (2020). Equipping teachers with globally competent practices: A mixed methods study on integrating global competence and teacher education. *International Journal of Educational Research, 103,* 101629–101629. https://doi.org/10.1016/j.ijer.2020.101629

Koran, J., & Jaffari, F. (2020). Deletion statistic accuracy in confirmatory factor models. *Methodological Innovations*, *13*(2). https://doi.org/10.1177/2059799120918349

Larsen, M., & Searle, M. (2017). International service learning and critical global citizenship: A cross-case study of a Canadian teacher education alternative practicum. *Teaching and Teacher Education*, *63*, 196–205. https://doi.org/10.1016/j.tate.2016.12.011

Malewski, E., & Phillion, J. (2009). International field experiences: The impact of class, gender and race on the perceptions and experiences of pre-service teachers. *Teaching and Teacher Education: An International Journal of Research and Studies, 25*(1), 52–60. http://dx.doi.org/10.1016/j.tate.2008.06.007

Mansilla, V. B., & Jackson, A. (2011). *Educating for global competency*. Asia Society. http://asiasociety.org/files/book-globalcompetence.pdf

Martorana F., Rania, N., & Lagomarsino, F. (2021). Which intercultural competences for teachers, educators, and social workers? A literature review. *International Journal of Intercultural Relations, 85*, 92–103. https://doi.org/10.1016/j.ijintrel.2021.09.003

Morais, D. B., & Ogden, A. C. (2011). Initial development and validation of the global citizenship scale. *Journal of Studies in International Education, 15*(5), 445–466. https://doi.org/10.1177/1028315310375308

Müller, F., Denk, A., Lubaway, E., Sälzer, C., Kozina, A., Perše, T., Rasmusson, M., Jugović, I., Nielsen, B., Rozman, M., Ojsteršek, A., & Jurko, S. (2020). Assessing social, emotional, and intercultural competences of students and school staff: A systematic literature review. *Educational Research Review*, *29*. https://doi.org/10.1016/j.edurev.2019.100304

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling, 9*(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8

National Center for Education Statistics. (2021). Racial/ethnic enrollment in public schools. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved April 8, 2022, from https://nces.ed.gov/programs/coe/indicator/cge.

Nielsen, B., Laursen, H., Reol, L., Jensen, H., Kozina, A., Vidmar, M., Rasmusson, M., Denk, A., Roczen, N., Jurko, S., Ojstersek, A., & Marusi, I. (2019). Social, emotional and intercultural competencies. A literature review with a particular focus on the school staff. *European Journal of Teacher Education*, *42*(3), 410–428. https://doi.org/10.1080/02619768.2019.1604670

O'Connor, K., & Zeichner, K. (2011). Preparing US teachers for critical global education. *Globalization, Societies and Education*, *9*(3-4), 521–536. https://doi.org/10.1080/14767724.2011.605333

OECD. (2018). *Preparing our Youth for an Inclusive and Sustainable World. The OECD PISA Global Competence Framework.* Paris: OECD. https://www.oecd.org/education/Global-competency-for-an-inclusive-world.pdf

OECD. (2019). *PISA 2018 Assessment and Analytical Framework*, Paris: OECD. https://doi.org/10.1787/b25efab8-en.

OECD. (2020). *PISA 2018 results (volume VI): are students ready to thrive in an interconnected world?.* Paris: OECD. https://doi.org/10.1787/d5f68679-en

Okken, G., Jansen, E., Hofman, W., & Coelen, R. (2019). Beyond the "welcome-back party": The enriched repertoire of professional teacher behaviour as a result of study abroad. *Teaching and Teacher Education, 86*, 102927–. https://doi.org/10.1016/j.tate.2019.102927

Oxley, L., & Morris, P. (2013). Global citizenship: A typology for distinguishing its multiple conceptions. *British Journal of Educational Studies, 61*(3), 301–325. https://doi.org/10.1080/00071005.2013.798393

Parkhouse, H., Tichnor-Wagner, A., Cain, J., & Glazier, J. (2016). "You don't have to travel the world": Accumulating experiences on the path toward globally competent teaching. *Teaching Education, 27*(3), 267–285. https://doi.org/10.1080/10476210.2015.1118032

Pashby, K., da Costa, M., Stein, S., & Andreotti, V. (2020). A meta-review of typologies of global citizenship education. *Comparative Education, 56*(2), 144–164. https://doi.org/10.1080/03050068.2020.1723352

Patterson, T. (2015). The transformative power of travel? Four social studies teachers reflect on their international professional development. *Theory and Research in Social Education, 43*(3), 345–371. https://doi.org/10.1080/00933104.2015.1064504

Perry, L. B., & Southwell, L. (2011). Developing intercultural understanding and skills: models and approaches. *Intercultural Education, 22*(6), 453–466. https://doi.org/10.1080/14675986.2011.644948

Portera, A. (2014). Intercultural competence in education, counselling and psychotherapy. *Intercultural Education (London, England), 25*(2), 157–174. doi.org/10.1080/14675986.2014.894176

Putman, & Byker, E. J. (2020). Global citizenship 1-2-3: Learn, think, and act. *Kappa Delta Pi Record, 56*(1), 16–21. https://doi.org/10.1080/00228958.2020.1696088

Robertson, S. L. (2021). Provincializing the OECD-PISA global competences project. *Globalisation, Societies and Education*, *19*(2), 167–182. https://doi.org/10.1080/14767724.2021.1887725

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Schleicher, A. (2015*). Schools for 21st-century learners: Strong leaders, confident teachers, innovative approaches.* OECD publishing. http://dx.doi.org/10.1787/23127090.

Shultz, L. (2007). Educating for global citizenship: Conflicting agendas and understandings. *Alberta Journal of Educational Research*, *53*(3), 248.

Siwatu, K. O., Putman, S. M., Starker-Glass, T. V., & Lewis, C. W. (2017). The culturally responsive classroom management self-efficacy scale: development and initial validation. *Urban Education*, *2*(7), 862–888. https://doi.org/10.1177/0042085915602534

Smolcic, E., & Katunich, J. (2017). Teachers crossing borders: A review of the research into cultural immersion field experience for teachers. *Teaching and Teacher Education, 62*, 47–59. https://doi.org/10.1016/j.tate.2016.11.002

Streiner, D. L. (1994). Figuring out factors: the use and misuse of factor analysis. *The Canadian Journal of Psychiatry, 39*(3), 135–140. https://doi.org/10.1177/070674379403900303.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson/ Allyn & Bacon.

Taie, S., & Goldring, R. (2020). Characteristics of public and private elementary and secondary school teachers in the United States: Results from the 2017–18 *National teacher and principal survey. First look* (NCES 2020-142). U.S. Department of Education. Washington, DC: National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020142.

Tichnor-Wagner, A., Parkhouse, H., Glazier, J., & Cain, J. M. (2016). Expanding approaches to teaching for diversity and social justice in K-12 education: Fostering global citizenship across the content areas. *Education Policy Analysis Archives, 24*(59), 1-31. https://doi.org/10.14507/epaa.24.2138

Trilokekar, R. D., & Kukar, P. (2011). Disorienting experiences during study abroad: Reflections of pre-service teacher candidates. *Teaching and Teacher Education: An International Journal of Research and Studies, 27*(7), 1141–1150. http://dx.doi.org/10.1016/j.tate.2011.06.002

van Werven, I. M., Coelen, R. J., Jansen, E. P. W. ., & Hofman, W. H. . (2023). Global teaching competencies in primary education. *Compare: A Journal of Comparative and International Education, 53*(1), 37-54. https://doi.org/10.1080/03057925.2020.1869520

VanderWeele, T. J., & Mathur, M. B. (2019). Some desirable properties of the Bonferroni correction: it the Bonferroni correction really so bad? *American Journal of Epidemiology, 188*(3), 617–618. https://doi.org/10.1093/aje/kwy250

Wahlström, N. (2014). Toward a conceptual framework for understanding cosmopolitanism on the ground. *Curriculum Inquiry, 44*(1), 113–132. https://doi.org/10.1111/curi.12038

Yong, A.G. & Pearce, S. (2013). A beginner's guide to factor analysis: focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology, 9*(2), 79–94. https://doi.org/10.1057/fsm.2014.17.

Zhao, Y. (2010). Preparing globally competent teachers: A new imperative for teacher education. *Journal of Teacher Education*, *61*(5), 422–431. https://doi.org/10.1177/0022487110375802

Zhou, Y. (2019). A mixed methods model of scale development and validation analysis. Measurement*, 17*(1), 38–47. https://doi.org/10.1080/15366367.2018.1479088

APPENDIX A: ASSESSMENT INSTRUMENT STUDY RECRUITMENT EMAIL - SCHOOL

This email will be sent to Global Education Administrators at Global Education Benchmark GROUP (GEBG) Schools and School Administrators at International Schools

Greetings,

My name is David Lynn and I serve as Director of International Studies at Charlotte Country Day School in Charlotte, North Carolina. In my role as a part-time doctoral candidate at the University of North Carolina-Charlotte, I am researching K-12 educator intercultural competencies growth. Over the past two years, I have worked on developing an assessment instrument aimed at measuring K-12 educator intercultural competencies and am currently in the process of validating this scale. Dr. Jae Hoon Lim, Professor of Educational Research, and Dr. Stella Kim, Assistant Professor of Educational Research, both at the University of North Carolina-Charlotte, are the faculty co-advisors on this research.

As a global educator and GEBG member, I am writing to invite you to participate in this current research study.

Why is this study being conducted?

The purpose of this study is to develop and validate an assessment instrument that aims to measure K-12 educator intercultural competencies. While intercultural scales do exist, few are specifically designed for K-12 educators. This research intends to investigate the validity of a scale based on intercultural competencies (attitudes, skills, knowledge, and action) and four theoretical factors, including global curriculum design, intercultural learning environment, student engagement, and personal intercultural competencies growth. The objective of this assessment instrument is to promote critical reflection and help guide professional development.

What is the criteria to participate?

Participation in this study is open to all K-12 educators who work in school settings. Staff and administrators are welcome to complete the assessment, but their results will not be included in the study.

What happens if I choose to participate?

If your school volunteers to participate in the study, you will be asked to distribute a link to the assessment to members of your K-12 faculty to complete between May 17 and June 10. The online assessment includes 30 multiple-choice items related to intercultural competencies and nine demographic questions. The assessment is designed to take 7 to 10 minutes to complete. All participant and individual school data will be kept confidential and identifiers will be removed from any published results. All participant schools will receive a summary of the results. Schools with 30 or more participants will receive an overview of their school’s responses to the 30 multiple-choice items to compare with data from all participating schools.

How do I participate?

If you are interested in volunteering as a research participant, please contact me at dlynn5@uncc.edu, indicating your interest. I will follow up with an email in mid-May and with a school-specific assessment link and introduction letter to send to your faculty. You are welcome to distribute this assessment during a regular faculty meeting or send the assessment link to your faculty to complete when their time allows before June 10.

Thank you for your participation in this study!

Sincerely,
David Lynn, Primary Investigator
david.lynn@charlottecountryday.org
dlynn5@uncc.edu

Dr. Jae Hoon Lim, Faculty Advisor
jhlim@uncc.edu

Dr. Stella Kim, Faculty Advisor
stella-kim@uncc.edu

APPENDIX B: ASSESSMENT INSTRUMENT STUDY RECRUITMENT EMAIL - FACULTY

This email will be provided to school representatives for distribution to potential faculty participants.

Greetings,

My name is David Lynn. In my role as an educator and part-time doctoral candidate at the University of North Carolina-Charlotte, I am researching K-12 educator intercultural competencies growth. Over the past two years, I have worked on developing an assessment instrument aimed at measuring K-12 educator intercultural competencies and am currently in the process of validating this scale. Dr. Jae Hoon Lim, Professor of Educational Research, and Dr. Stella Kim, Assistant Professor of Educational Research, both at the University of North Carolina-Charlotte, are the faculty advisors on this research.

As a fellow educator, I am writing to invite you to participate in the current research study.

Why is this study being conducted?

The purpose of this study is to develop and validate an assessment instrument that aims to measure K-12 educator intercultural competencies. While intercultural scales do exist, few are specifically designed for K-12 educators. This research intends to investigate the validity of a scale based on intercultural competencies (attitudes, skills, knowledge, and action) and four theoretical factors, including global curriculum design, intercultural learning environment, student engagement, and personal intercultural competencies growth. The objective of this assessment instrument is to promote critical reflection and help guide professional development.

What is the criteria to participate?

Participation in this study is open to all K-12 faculty who work in school settings. Staff and administrators are welcome to complete the assessment, but their results will not be included in the study.

What happens if I choose to participate?

If you volunteer for the study, you will be asked to complete an online assessment that includes 30 multiple choice items related to intercultural competencies and nine demographic questions. The assessment is designed to take 7 to 10 minutes to complete. All individual participant information will be kept anonymous and school identifiers removed, in order to ensure complete confidentiality.

How do I participate?

Please click on the following link to the Google Form assessment: (CLICK HERE)

Please complete the form by Friday, June 10. If you have any questions or are interested in learning more, please contact me at dlynn5@uncc.edu

Thank you for your participation in this study!

Sincerely,
David Lynn, Primary Investigator
david.lynn@charlottecountryday.org
dlynn5@uncc.edu

Dr. Jae Hoon Lim, Faculty Advisor
jhlim@uncc.edu

Dr. Stella Kim, Faculty Advisor
stella-kim@uncc.edu

APPENDIX C:

Final Intercultural Competency Measure for Educators (ICME) Items

Item		IC Domain
Curriculum		
C.5	My knowledge of global issues and topics informs the content I teach.	Knowledge
C.6	My understanding of global perspectives is reflected in lessons I teach.	Knowledge
C.2	I believe that reaching second language proficiency is a priority.	Attitudes
C.4	I adapt my curriculum content to include diverse global perspectives.	Skills
C.3	I evaluate student global learning as it relates to my content area.	Skills
Diverse Student Inclusion		
L.3	I create a classroom where culturally diverse students feel valued.	Skills
S.7	I advocate for students whose home culture differs from societal norms.	Action
L.5	I know strategies for building inclusive global learning environments.	Knowledge
S.6	I ensure culturally diverse students have a voice.	Action
L.6	I ensure cultural diversity is clearly integrated into classroom settings.	Action
Cross-Cultural Openness		
S.2	I am open to new communication strategies that engage culturally diverse students.	Attitudes
S.1	I am eager to learn about the cultures represented by my students.	Attitudes
L.2	I'm curious how diverse student cultures can be better represented.	Attitudes
C.1	I value the inclusion of global content across the curriculum.	Attitudes
P.1	I enjoy exploring unfamiliar cultures and people by reading books, watching films, or sampling unique ethnic foods.	Attitudes
Collaboration and Adaptation		
C.7	I discuss strategies for implementing global content with colleagues.	Action
L.7	I collaborate with colleagues in promoting a culturally inclusive school.	Action
C.8	I seek out and break down cultural stereotypes in curriculum.	Action
P.3	I adapt my behavior and mannerisms when engaging cross-culturally.	Skills
S.3	I adapt pedagogy to engage culturally diverse students.	Skills
Systematic Awareness		
S.4	I am aware of ways cultural diversity impacts student engagement.	Knowledge
P.6	I know how existing power structures impact culturally diverse people.	Knowledge
S.5	I recognize how systemic inequities inhibit the inclusion of culturally diverse learners.	Knowledge
P.7	I actively support people marginalized due to cultural differences.	Action
L.4	I identify cultural hierarchies that influence learning environments.	Skills

Note. All items will be answered using a six-point Likert Scale: (a) Strongly Disagree, (b) Disagree, (c) Slightly Disagree, (d) Slightly Agree, (e) Agree, and (f) Strongly Agree